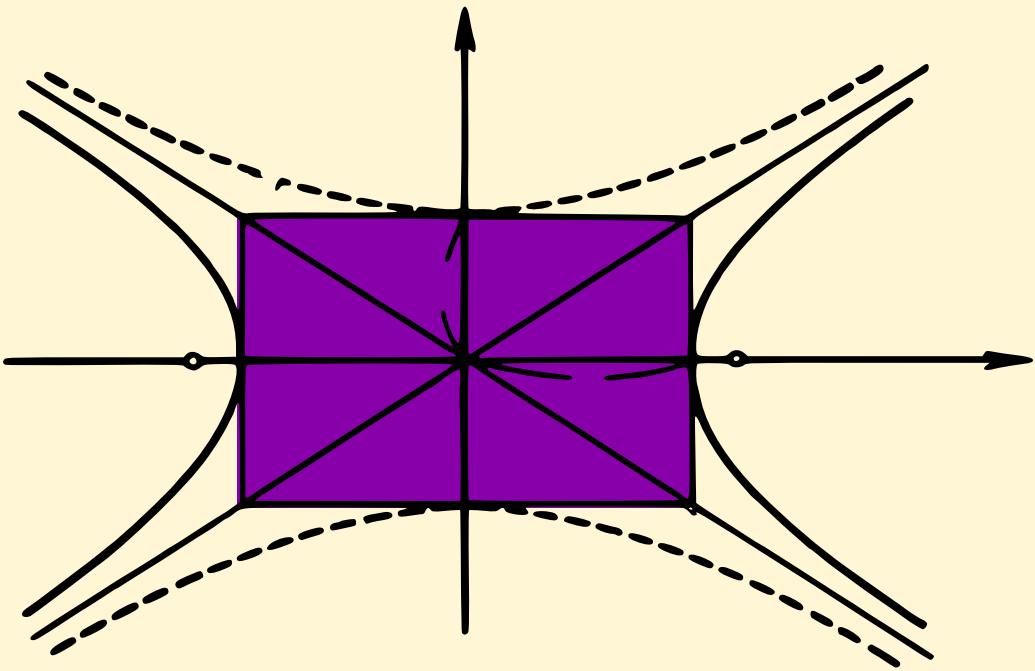
V. Chipatchev

MATHÉ-MATIQUES SUPÉRIEURES



Éditions Mir Moscou

В. С. ШИПАЧЕВ

ВЫСШАЯ МАТЕМАТИКА

Под редакцией
академика А. Н. Тихонова

ИЗДАТЕЛЬСТВО «ВЫСШАЯ ШКОЛА»
МОСКВА

V. CHIPATCHEV

MATHÉMATIQUES SUPÉRIEURES

Sous la direction de l'académicien
A. TYKHONOV

ÉDITIONS MIR MOSCOU

Traduit du russe
par Djilali EMBAREK

На французском языке

Imprimé en Union Soviétique

ISBN 5—03—000006—2

© Издательство « Высшая школа », 1985
© traduction française, Editions Mir, 1988

PRÉFACE

Le siècle du progrès technique et scientifique est marqué par la pénétration des méthodes mathématiques dans toutes les sphères des activités humaines. Les sciences connexes se servent des méthodes mathématiques pour poser de nouveaux problèmes qui, par contrecoup, favorisent le développement de ces dernières. Les mathématiques sont devenues une discipline indispensable même aux non-mathématiciens.

L'étude des mathématiques développe la logique, fournit les éléments nécessaires à l'appréhension des problèmes complexes auxquels est confrontée l'économie nationale.

Le présent ouvrage a été écrit sur la base du manuel *Cours de mathématiques supérieures* publié aux Editions de l'Université Lomonossov de Moscou. Cet ouvrage a valu à son auteur la médaille de bronze de l'Exposition des réalisations de l'Economie Nationale d'U.R.S.S. en 1983.

L'auteur a lu ce cours pendant de longues années à la faculté de géologie de l'Université de Moscou. Il a en même temps enseigné aux Cours nationaux de recyclage des professeurs de mathématiques du secondaire qui sont organisés sous l'égide de l'Université de Moscou. La longue expérience pédagogique de l'auteur s'est concrétisée aussi bien dans le choix des sujets abordés que dans le caractère de l'exposé. La concision de l'ouvrage ne sacrifie ni à la clarté ni à la rigueur.

Les nombreux exemples et exercices qui illustrent les notions et théorèmes fondamentaux aideront les élèves à assimiler plus profondément les bases des mathématiques et les guideront dans la résolution de problèmes pratiques.

L'heureuse conjugaison de la simplicité et de la rigueur, le choix judicieux des exemples permettent d'utiliser cet ouvrage comme manuel pour les disciplines non mathématiques de l'enseignement supérieur.

André Tykhonov

AVANT-PROPOS

Cet ouvrage expose les bases des mathématiques supérieures et à ce titre peut être utilisé aussi bien dans les universités que dans les écoles techniques supérieures où les diverses branches des mathématiques supérieures sont regroupées dans un même cours.

L'auteur s'est efforcé dans la mesure du possible de faire un exposé exhaustif, rigoureux et clair dont l'objectif avoué n'est pas seulement de communiquer telle ou telle information sur les mathématiques, mais d'aiguiser l'intérêt du lecteur pour cette discipline, d'élargir son horizon et le lui inculquer une certaine culture mathématique.

Cet ouvrage se réclame du principe didactique : du simple au complexe. Ainsi la notion de limite est d'abord étudiée pour les suites numériques, puis pour les fonctions d'une variable. Sont ensuite introduites la notion de limite des sommes intégrales et, enfin, des fonctions de plusieurs variables.

Les nombres réels sont définis par une méthode axiomatique qui permet d'exposer de la manière la plus concise les généralités nécessaires sur les nombres et de passer plus aisément à l'exposition directe du matériel.

L'analyse mathématique — calcul différentiel et calcul intégral — constitue la charnière de ce manuel. La notion de fonction, qui est capitale en mathématiques, est abordée au secondaire, mais ce n'est qu'en mathématiques supérieures qu'elle est l'objet d'une étude approfondie et systématique. Signalons que cette notion est définie au moyen de la notion d'ensemble, ce qui correspond à l'esprit moderne des mathématiques.

L'auteur exprime sa profonde gratitude à l'académicien A. Tykhonov, aux professeurs A. Babaev, V. Boutouzov, A. Chestakov, M. Govorov, D. Kostomarov, N. Matvéev, A. Svechnikov, G. Tchandirov pour leur précieux concours à la parution de cet ouvrage.

Victor Chipatchev

INTRODUCTION

Les mathématiques sont à la fois la plus antique et la plus jeune des sciences. Elles ont commencé à se former au deuxième millénaire avant J.-C. sous l'impulsion du commerce, de l'agriculture et de la navigation qui nécessitaient des méthodes de calcul et de mesure qui plongent leurs racines dans la plus haute antiquité. Les bâtisseurs des pyramides égyptiennes connaissaient déjà les mathématiques.

Dans la Grèce Antique, les mathématiques s'érigent en science autonome à partir du VI^{e} siècle avant J.-C. Elles prennent définitivement corps au III^{e} siècle dans les immortels *Eléments* d'Euclide. Ce monumental ouvrage a servi de base à l'enseignement de la géométrie pendant plus de deux mille ans. Ceci différencie fondamentalement les mathématiques des autres sciences. Les concepts physiques d'Aristote qui nous semblent quelque peu simplistes sont devenus l'apanage de l'histoire des sciences, bien qu'ils aient généralisé toutes les connaissances accumulées à l'époque sur le monde ambiant. Or, le théorème de Pythagore continue d'être un pilier de la géométrie.

Les mathématiques ne cessent d'évoluer, d'élaborer de nouvelles méthodes, de bourgeonner, de perfectionner leur symbolique et leurs outils. L'émergence de la physique des Temps modernes est redevable à l'application des mathématiques par Kepler et Galilée à l'étude des phénomènes terrestres et célestes. Au XVI^{e} siècle, la géométrie analytique de Descartes et le calcul différentiel et intégral de Newton et de Leibniz constituent un tournant décisif dans l'histoire des mathématiques et marquent fortement aussi bien les mathématiques proprement dites que les autres sciences telles la physique et l'astronomie.

L'intense développement qui s'est ensuivi a conduit à la charnière du XIX^{e} et du XX^{e} siècle à une nouvelle révolution scientifique liée, en particulier, à la reconnaissance du droit de cité des géométries non euclidiennes (géométries de Lobatchevski, Riemann, Bolyai) et à la création de la théorie des ensembles par Cantor. Les mathématiques poursuivent leur marche triomphale en forçant l'admiration par la diversité des branches spéciales,

la nouveauté et l'originalité des concepts et méthodes, les particularités du langage. Le processus de différenciation des sciences a envahi les mathématiques en les enrichissant d'une multitude de domaines.

Le développement des méthodes et des branches des mathématiques s'est accompagné de leur implantation dans les autres sciences, d'un processus de mathématisation des sciences. De par la logique du développement des sciences, les mathématiques se sont transformées en une méthode de recherche scientifique. Si à l'époque de la physique classique les mathématiques servaient essentiellement au traitement des données expérimentales, à l'établissement d'une relation exacte entre les phénomènes et processus physiques, à la fin du XIX^{e} siècle les calculs mathématiques suggéraient des hypothèses physiques et des découvertes. Les mathématiques ne servaient pas uniquement à traiter les indications des instruments de mesure et les résultats des expériences, mais aussi à élaborer des modèles mathématiques dont la véritable signification physique restait à élucider.

En clair, il s'agit d'utiliser les méthodes mathématiques pour pénétrer dans des domaines du monde physique encore vierges et provisoirement inaccessibles par les méthodes physiques, à y découvrir des lois mathématiques, à élaborer des modèles mathématiques décrivant des processus physiques inconnus et, partant, à orienter la pensée de l'expérimentateur. Aujourd'hui le physicien théoricien est avant tout un mathématicien. « Les mathématiques, a dit l'éminent savant américain F. J. Dyson, ne sont pas pour le physicien seulement un instrument à l'aide duquel il peut décrire quantitativement tout phénomène, mais la principale source d'idées et de principes générateurs de nouvelles théories ».

Un exemple remarquable du rôle du raisonnement mathématique dans les découvertes physiques est fourni par la relativité généralisée d'Einstein dont la construction a été achevée avant toute expérimentation et dont les hypothèses ont été brillamment confirmées par les résultats.

Un exemple non moins significatif est l'apparition de la mécanique quantique qui a été élaborée mathématiquement à partir de données physiques connues mais inexplicables, par un trait génial de l'imagination mathématique. Cette théorie a non seulement été confirmée par les expériences, mais elle est devenue la source et le moteur de l'essor de la physique atomique avec les résultats impressionnants que l'on connaît.

Tous ceux qu'ils soient mathématiciens, physiciens ou philosophes qui se sont intéressés de près à l'histoire des mathématiques et à leurs applications dans les sciences naturelles et techniques et qui ont médité sur le rapport des mathématiques avec le monde objectif se sont inévitablement interrogés sur la merveilleuse capacité des mathématiques de décrire correctement ou de refléter les

processus physiques ou le comportement de l'Univers physique. Les précurseurs de la science, tels Kepler et Galilée, sous le coup des résultats acquis, s'imaginaient que le livre de la nature a été écrit par son Créateur divin dans le langage mathématique et qu'il ne restait au savant qu'à le lire. Depuis une foule d'hypothèses ont été avancées sur la nature des mathématiques et leurs vertus cognitives sans qu'aucune prévaille.

Au XX^{e} siècle le philosophe autrichien Wittgenstein a émis l'étonnante idée que les mathématiques ne sont rien d'autre qu'un ensemble de tautologies et les démonstrations mathématiques, des transformations tautologiques. Cette théorie expliquait l'authenticité absolue des mathématiques et de leur applicabilité universelle. Mais elle était incapable de trouver une explication à la faculté des mathématiques de découvrir des faits nouveaux dans le monde, faculté qui est primordiale au développement de la science et qui permet d'appliquer largement les mathématiques aux sciences spéciales.

Il importe de souligner que les mathématiques procèdent par des concepts bien plus abstraits que les autres sciences. Même le nombre naturel le plus ordinaire, par exemple 4, fait abstraction des traits spécifiques des objets décrits (arbres, pieds d'une table, coins d'une salle, etc.) et ne caractérise qu'une classe de quatre éléments. La notion même de nombre naturel est un concept bien plus abstrait encore, car elle représente la classe de toutes les classes contenant plus d'un élément.

La force des mathématiques est justement dans leur capacité de créer des concepts de plus en plus abstraits, d'opérer sur eux pour étudier leurs particularités et les lois qui les régissent. C'est précisément pour cette raison que les méthodes mathématiques peuvent être utilisées dans les sciences autres que la physique à mesure que ces sciences deviennent théoriques, c'est-à-dire qu'elles se dotent de concepts suffisamment abstraits qu'elles appliqueront à leur développement.

Le philosophe allemand Kant a dit qu'une science mérite d'autant plus son nom qu'elle contient de mathématiques. A l'époque, les mathématiques n'étaient partie intégrante que de la mécanique, la physique et l'astronomie. De nos jours le niveau théorique des sciences s'est tellement élevé que leur fusion est devenue non seulement possible, mais absolument nécessaire pour le développement tant de ces sciences que des mathématiques elles-mêmes.

Il va de soi que le processus de mathématisation n'a pas touché toutes les sciences au même degré. Une réussite est l'application des méthodes mathématiques aux sciences minérales et en biologie. Ceci a été possible grâce à la pénétration de la biologie dans les processus intracellulaires et à leur analyse à un niveau moléculaire. A titre d'exemple on pourrait citer les études du fonctionnement et la conception de modèles de certaines fonctions du neuro-

ne, l'étude de problèmes d'hérédité, etc.

Même dans les sciences sociales qui sont le plus longtemps restées réfractaires aux mathématiques, excepté les méthodes statistiques d'analyse de certains processus et phénomènes sociaux, on pourrait citer aussi de nombreux domaines dans lesquels les mathématiques ont donné de bons résultats : problèmes de démographie, de linguistique structuraliste, etc. Mais le succès le plus spectaculaire est l'implantation des méthodes mathématiques en économie et dans la gestion des processus économiques. La gestion scientifique de tels processus dans une économie planifiée ne peut aujourd'hui être réalisée que sur la base de méthodes mathématiques exactes dans tous les domaines — prospection des gisements, étude du marché de l'offre et de la demande, étude de la main-d'œuvre, transport de voyageurs et de marchandises, intervention sur les phénomènes atmosphériques — bref, la vie contemporaine est impossible sans les mathématiques.

Mais de quelles mathématiques est-il question ici : des mathématiques pures ou des mathématiques appliquées ? Cette démarcation traditionnelle est aujourd'hui de plus en plus formelle et se vide progressivement de son sens. Les sections mêmes les plus abstraites des mathématiques pures peuvent semble-t-il être matérialisées dans les domaines les plus inattendus des sciences et des techniques. En même temps la résolution des problèmes pratiques et théoriques stimule l'élaboration de nouvelles méthodes abstraites et engendre de nouvelles branches mathématiques.

Les dernières décennies ont été marquées par un intense développement des instruments et méthodes du calcul numérique. La simulation mathématique permet d'étudier des processus qui sont inexpérimentables (fusion thermonucléaire contrôlée, physique du plasma, lasers, etc.). « De fait, une nouvelle orientation prévaut dans les recherches physiques théoriques des deux dernières décennies », affirme l'académicien soviétique A. Samarski. « L'ordinateur permet grâce au modèle mathématique d'étudier les structures et les processus physiques, de simuler leur comportement dans des conditions différentes, de trouver les paramètres et régimes optimaux des constructions réelles et théoriques ». Des horizons tout à fait nouveaux s'ouvrent aux mathématiques.

L'ère de la révolution scientifique et technique est l'ère de la mathématisation de la science, de la technique, de l'économie et de la gestion. Tout chercheur scientifique ou ingénieur se doit non seulement de connaître les bases des mathématiques, mais aussi de bien maîtriser les toutes dernières méthodes mathématiques utilisables dans sa spécialité. Aujourd'hui tout travail scientifique sérieux doit se prévaloir des mathématiques. On peut en toute conscience dire que l'étude des mathématiques contribue à former la pensée scientifique et sa large utilisation est un gage de progrès dans les sciences et techniques.

PREMIÈRE PARTIE

ANALYSE MATHÉMATIQUE DES FONCTIONS D'UNE VARIABLE

CHAPITRE PREMIER

NOMBRES RÉELS

§ 1. Ensembles. Notations. Symboles logiques

La notion d'ensemble est fondamentale en mathématiques. C'est une notion première qui est indéfinissable par d'autres notions. Les termes « groupement », « famille », « système », « collection », etc., sont synonymes d'« ensemble ». Comme exemples d'ensembles citons l'ensemble des étudiants dans une salle de cours ; la famille des étoiles de la Grande Ourse ; un système de trois équations à trois inconnues ; l'ensemble des entiers, etc. On voit sur ces exemples qu'un ensemble peut être constitué d'un nombre fini ou infini d'objets.

Les objets qui composent un ensemble s'appellent *éléments* ou *points* de cet ensemble. Les ensembles sont représentés par des lettres majuscules, leurs éléments, par des lettres minuscules. Si x est un élément d'un ensemble X, on note $x \in X$ et on lit x appartient à X. Si x n'est pas élément de X, on note $x \notin X$ (x n'appartient pas à X). La notation $X = \{x_1, \ldots, x_n\}$ exprime que X est composé des éléments $x_1, \ldots, x_n$. *Idem* pour la notation $X = \{x_1, x_2, x_3, \ldots\}$.

Soient X et Y deux ensembles. Si X et Y sont composés des mêmes éléments, on dit qu'ils sont confondus et l'on écrit $X = Y$. Si tous les éléments de X appartiennent à Y, on dit que X est contenu, ou inclus, dans Y ou encore que X est un sous-ensemble ou une partie de Y. On note ceci $X \subset Y$ ou $Y \supset X$ (Y contient X). Si X n'est pas inclus dans Y, on écrit $X \not\subset Y$. L'ensemble vide est d'un usage fréquent en mathématiques. Cet ensemble ne contient aucun élément et se représente par le symbole $\varnothing$. L'ensemble vide est une partie de tout ensemble.

Dans la suite nous aurons affaire à divers ensembles de nombres réels. Pour abréger on omettra le qualificatif « réel » si aucune confusion n'est à craindre.

Soit $P(x)$ une propriété d'un nombre x. L'écriture $\{x \mid P(x)\}$ désigne l'ensemble de tous les nombres possédant la propriété $P(x)$. Ainsi $\{x \mid x^2 - 3x + 2 = 0\}$ est l'ensemble des zéros de l'équation $x^2 - 3x + 2 = 0$: il est constitué des éléments 1 et 2 ; $\{x \mid 3 < x < 7\}$ est l'ensemble des nombres satisfaisant la double inégalité $3 < x < 7$; $\{x \mid x > 7 \text{ et } x < 3\} = \varnothing$.

Si $x_1, \ldots, x_n$ sont des nombres arbitraires, la notation $x = \max\{x_1, \ldots, x_n\}$ (resp. $x = \min\{x_1, \ldots x_n\}$) exprime que x est le maximum (resp. le minimum) des nombres $x_1, \ldots, x_n$.

Certains termes et expressions se répètent souvent dans les définitions, théorèmes, etc. Aussi avons-nous intérêt à utiliser en pareilles occasions les symboles logiques qui sont économiques.

Nous n'indiquerons que les symboles logiques les plus simples et les plus usités : le quantificateur existentiel $\exists$ et le quantificateur universel $\forall$. Le symbole $\exists$ se lit « il existe », le symbole $\forall$, « quel que soit » ou « pour tout ». Ainsi $\exists\, x \in X \ldots$ signifie : « il existe un nombre x de X tel que $\ldots$ » ; $\forall x \in X$: « pour tout nombre x de X » ; $\forall x \in X : \alpha$ exprime que « pour tout nombre x de X on a (ou est réalisée) la proposition α ». Par exemple, l'expression

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \neq x_0,\ |x - x_0| < \delta : |f(x) - A| < \varepsilon$$

se lit : « quel que soit $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout x distinct de x_0 et satisfaisant l'inégalité $|x - x_0| < \delta$ l'on a $|f(x) - A| < \varepsilon$ ».

Le symbole ■ exprimera la fin d'une démonstration.

§ 2. Les nombres réels et leurs propriétés fondamentales

Du cours de mathématiques élémentaires on sait que l'ensemble des réels est composé des nombres rationnels et irrationnels. Un nombre rationnel est un nombre qui se représente par un quotient d'entiers p et q ($q \neq 0$). Un nombre réel qui n'est pas rationnel est irrationnel. Tout nombre rationnel est soit entier, soit se représente par un développement décimal limité ou périodique illimité. Un nombre irrationnel se représente, lui, par un développement décimal non périodique illimité. Ainsi les rationnels 3/4 et 1/3 ont pour développements décimaux 0,75 et 0,333... ; les nombres irrationnels $\sqrt{2}$ et π ont des développements décimaux illimités : $\sqrt{2} = 1{,}41421356\ldots$, $\pi = 3{,}14159\ldots$

Récapitulons nos connaissances sur les réels, énumérons leurs propriétés fondamentales et déduisons ensuite quelques corollaires.

I. Addition et multiplication des réels

A tout couple de nombres réels a et b sont associés de façon unique deux nombres réels $a + b$ et $a \cdot b$ appelés respectivement leur *somme* et leur *produit* et vérifiant les propriétés suivantes. Quels que soient les nombres a, b et c :

1° $a + b = b + a$ (*commutativité*).

2° $a + (b + c) = (a + b) + c$ (*associativité*).

3° $a \cdot b = b \cdot a$ (*commutativité*).

4° $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (*associativité*).

5° $(a + b) \cdot c = ac + bc$ (*distributivité pour l'addition*).

6° Il existe un seul élément neutre 0 pour l'addition tel que $a + 0 = a$ quel que soit a.

7° Pour tout nombre a il existe un nombre $(-a)$ dit opposé de a tel que $a + (-a) = 0$.

8° Il existe un seul élément neutre 1 pour la multiplication tel que pour tout a on a $a \cdot 1 = a$.

9° Pour tout nombre $a \neq 0$ il existe un nombre a^{-1}, appelé *inverse de* a, tel que $a \cdot a^{-1} = 1$; le nombre a^{-1} est noté aussi par $\frac{1}{a}$.

II. Comparaison des nombres réels

Entre deux nombres réels quelconques a et b on a l'une des relations suivantes : $a = b$ (a égal b), $a > b$ (a strictement supérieur à b) ou $b > a$. La relation $=$ est transitive : $[a = b$ et $b = c] \Rightarrow [a = c]$.

La relation $>$ jouit des propriétés suivantes. Quels que soient les nombres a, b et c :

10° $[a > b$ et $b > c] \Rightarrow [a > c]$ (*transitivité*).

11° $[a > b] \Rightarrow [a + c > b + c]$ (*régularité*).

12° $[a > 0$ et $b > 0] \Rightarrow [a \cdot b > 0]$.

La relation $a > b$ peut encore s'écrire $b < a$ (b strictement inférieur à a). La relation $a \geqslant b$ (ou, ce qui est équivalent, $b \leqslant a$) exprime que a est soit égal, soit strictement supérieur à b. Les relations $a < b$, $a > b$, $a \leqslant b$, $a \geqslant b$ sont des *inégalités*, les deux premières sont *strictes*, les deux dernières, *larges*.

III. Continuité des nombres réels

13° Soient X et Y deux ensembles de réels. Si quels que soient $x \in X$ et $y \in Y$, on a $x \leqslant y$, il existe alors au moins un c tel que $x \leqslant c \leqslant y$.

A noter que la propriété de continuité est caractéristique de l'ensemble des réels mais pas de l'ensemble des rationnels. En effet, soient X l'ensemble des rationnels x tels que $x < \sqrt{2}$ et Y l'ensemble des rationnels y tels que $y > \sqrt{2}$. Il est évident que quels que soient $x \in X$ et $y \in Y$ on a $x \leqslant y$, or il n'existe pas de rationnel c tel que $x \leqslant c \leqslant y$. En effet, seul le nombre $\sqrt{2}$ qui, on le sait, est irrationnel jouit de cette propriété.

Les autres propriétés des nombres réels résultent des propriétés I, II et III. Penchons-nous sur quelques-unes d'entre elles, les autres seront utilisées dans la suite sans démonstration formelle.

Quels que soient les nombres a, b, c et d :

14° Le nombre $x = b + (-a)$ est solution (ou racine) de l'équation $a + x = b$.

En effet, d'après les propriétés 1°, 2°, 6° et 7° on a : $a + b + (-a) = b$. ■

Le nombre $b + (-a)$ s'appelle *différence* des nombres b et a et se note $b - a$. A noter que si $a < b$ (ou, ce qui revient au même, $b > a$), la différence $b - a > 0$. En effet, l'inégalité $b > a$ entraîne en vertu de 11° : $b + (-a) > a + (-a)$ ou $b - a > 0$.

15° Le nombre $x = ba^{-1}$ est solution (ou racine) de l'équation $ax = b$ si $a \neq 0$.

En effet, d'après les propriétés 3°, 4°, 8° et 9°, on a : $a \cdot ba^{-1} = b$. ■

Le nombre ba^{-1} s'appelle *quotient* des nombres b et a et se note $\frac{b}{a}$ ou $b : a$.

16° Si $a < b$, alors $-a > -b$.

En effet, puisque $a < b$, il vient $b - a > 0$. En vertu de 11° on a donc $b - a + (-b) > 0 + (-b)$, d'où $-a > -b$. ■

En particulier, si $a > 0$, alors $-a < 0$ et si $a < 0$, alors $-a > 0$ (on se sert du fait que $-0 = 0$; en effet, $(-0) + 0 = -0$ d'après 6° et $(-0) + 0 = 0$ d'après 7°, d'où $-0 = 0$).

17° Si $a > b$ et $c > d$, alors $a + c > b + d$.

En effet, si $a > b$ et $c > d$, en vertu de 11° on a $a + c > b + c$ et $c + b > d + b$. Donc $a + c > b + d$ d'après 10°. ■

18° Si $a < b$ et $c > d$, alors $a - c < b - d$.

En effet, puisque $c > d$, la propriété 16° nous donne $-c < -d$. En ajoutant terme à terme les inégalités $a < b$ et $-c < -d$ (ce qui est possible d'après 17°) on trouve $a - c < b - d$. ■

19° $a - a = 0$.

En effet, $a - a = a + (-a) = 0$. ■

20° $a \cdot 0 = 0$.

On a $a \cdot 0 = a \cdot (b - b) = ab - ab = 0$. ■

21° $-(-a) = a$.

On a $-(-a) = (-(-a)) + (-a) + a = 0 + a = a$. ■

22° $(-a)\, b = -ab$.

En effet, $(-a)\, b = (-a)\, b + ab + (-ab) = [(-a) + a] \cdot b - ab = 0 \cdot b - ab = 0 - ab = -ab$. ■

Signalons que la propriété 5° nous a permis de remplacer la somme $(-a)\, b + ab$ par le produit $[(-a) + a]\, b$. La propriété 22° nous donne en particulier : $(-1)a = -a$.

23° Si $a < 0$ et $b > 0$, alors $ab < 0$.

Puisque $a < 0$, il vient $-a > 0$. La propriété 12° nous donne $(-a)\, b > 0$. Donc $(-a)b = -ab > 0$ et par suite $ab < 0$. ■

24° Si $a < 0$ et $b < 0$, alors $ab > 0$.

En effet, puisque $b < 0$, on a $-b > 0$. D'après la propriété 23°, $(-b)\, a < 0$. Donc $(-b)\, a = -ab < 0$ et par suite $ab > 0$. ■

25° Si $a \neq 0$, alors $a \cdot a = a^2 > 0$.

Ceci résulte des propriétés 12° et 24°. En particulier, $1 = 1^2 > 0$, c'est-à-dire que $1 > 0$.

26° Si $a > 0$, alors $a^{-1} > 0$.

En effet, $aa^{-1} = 1 > 0$ d'après les propriétés 9° et 25°. Si l'on admet que $a^{-1} \leqslant 0$, les propriétés 20° et 23° nous disent que $aa^{-1} \leqslant 0$. Cette contradiction montre que $a^{-1} > 0$. ■

Nous voyons donc que les propriétés I, II et III des nombres réels entraînent toutes les autres. On peut donc dire que les nombres réels forment un ensemble d'éléments possédant les propriétés I, II et III. Une telle définition est dite *axiomatique*, quant aux propriétés I, II et III, elles sont appelées *axiomes* des nombres réels.

Signalons en conclusion que les propriétés I, II et III nous permettent de représenter tout nombre réel par un développement décimal illimité :

$$a,\ a_1a_2a_3 \ldots a_n \ldots$$

où a est un entier quelconque, a_1, a_2, a_3, . . ., a_n des entiers compris entre 0 et 9 ($0 \leqslant a_n \leqslant 9$). Nous glisserons cependant sur cette question.

§ 3. Représentation géométrique des nombres réels

1. Représentation des nombres réels par des points de la droite. Introduisons quelques notions préliminaires. Considérons une droite arbitraire. On peut la munir de deux orientations opposées. Fixons l'une de ces orientations (fig. 1). Choisissons par ail-

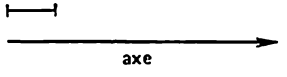


Fig. 1

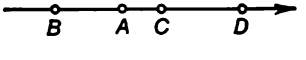


Fig. 2

leurs une unité pour mesurer la longueur des segments. Une droite sur laquelle on a choisi une orientation s'appelle *droite orientée*, ou *axe*.

Soient deux points quelconques A et B sur un axe. On dit que le segment AB est *orienté* si l'on indique lequel des points A et B est son origine et lequel, son extrémité. On désignera un segment orienté d'origine A et d'extrémité B par $\overrightarrow{AB}$ et on conviendra qu'il est orienté de A vers B. Signalons que dans l'écriture d'un segment orienté la première lettre représente l'origine, la seconde, l'extrémité. La longueur d'un segment orienté $\overrightarrow{AB}$ se désigne par $|\overrightarrow{AB}|$ ou AB.

La notion de mesure algébrique est introduite pour les segments orientés situés sur un même axe ou sur des axes parallèles. On appelle *mesure algébrique* $\overline{AB}$ d'un segment orienté $\overrightarrow{AB}$ le nombre algébrique égal à AB ou $-AB$ selon que $\overrightarrow{AB}$ est de mê-

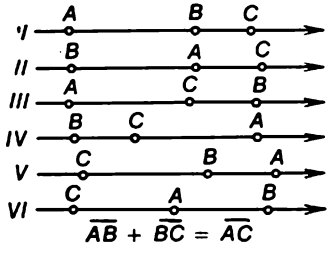


Fig. 3

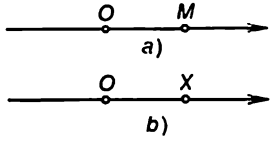


Fig. 4

me sens ou de sens contraire à l'axe. Dans le cas des segments $\overrightarrow{AB}$ et $\overrightarrow{CD}$ de la figure 2, on a $\overline{AB} = -AB$, $\overline{CD} = CD$.

Signalons que quelle que soit l'orientation de l'axe, les mesures algébriques des segments $\overrightarrow{AB}$ et $\overrightarrow{BA}$ sont de signes contraires :

$$\overline{AB} = -\overline{BA}.$$

Si les points A et B sont confondus, on convient que la mesure algébrique $\overline{AB}$ est nulle.

Pour trois points arbitraires A, B et C d'un axe on a la relation

$$\overline{AB} + \overline{BC} = \overline{AC}$$

dite *relation de Chasles*, qui sera fréquemment utilisée dans la suite.

On établit de façon élémentaire la relation de Chasles en envisageant une par une les six dispositions des points A, B et C sur l'axe (fig. 3).

Passons maintenant à la représentation géométrique des nombres réels. Appelons *axe de coordonnées* une droite sur laquelle on a fixé une orientation et un point O (l'origine des coordonnées). (On admet qu'on a fait choix d'une unité de mesure.) Soit M un point quelconque de l'axe (fig. 4, *a*). Associons au point M un nombre x égal à la mesure algébrique $\overline{OM}$ du segment $\overrightarrow{OM}$. Le nombre x s'appelle *coordonnée* du point M. A chaque point de l'axe de coordonnées on peut associer un nombre réel bien défini : sa coordonnée. La réciproque est vraie : à chaque nombre réel x correspond un point de l'axe de coordonnées, plus exactement le point M de coordonnée x.

Les nombres réels peuvent donc être représentés par les points de l'axe de coordonnées. C'est pourquoi les points de l'axe de coordonnées sont souvent suivis d'un nombre: leur coordonnée (fig. 4, *b*).

Soient M_1 et M_2 deux points de coordonnées respectives x_1 et x_2 (fig. 5). Exprimons la mesure algébrique $\overline{M_1M_2}$ en fonction des coordonnées des points M_1 et M_2. D'après la relation de Chasles

$$\overline{OM_1} + \overline{M_1M_2} = \overline{OM_2},$$

d'où $\overline{M_1M_2} = \overline{OM_2} - \overline{OM_1}$. Or $\overline{OM_1} = x_1$ et $\overline{OM_2} = x_2$, donc

$$\overline{M_1M_2} = x_2 - x_1.$$

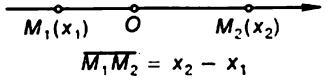


Fig. 5

Cette formule sera souvent utilisée en géométrie analytique.

2. Ensembles numériques d'usage fréquent. Soient $a < b$ deux nombres. On se servira des notations suivantes:

$\{x \mid a \leqslant x \leqslant b\} = [a, b]$; $\{x \mid a < x \leqslant b\} =]a, b]$;

$\{x \mid a \leqslant x < b\} = [a, b[$;

$\{x \mid a < x < b\} =]a, b[$;

$\{x \mid a \leqslant x\} = [a, +\infty[$; $\{x \mid a < x\} =]a, +\infty[$;

$\{x \mid x \leqslant b\} =]-\infty, b]$; $\{x \mid x < b\} =]-\infty, b[$.

L'ensemble des réels sera désigné comme suit:

$$\{x \mid -\infty < x < +\infty\} \text{ ou }]-\infty, +\infty[.$$

Tous ces ensembles s'appellent *intervalles*:

$[a, b]$ est un intervalle *fermé borné*, ou *segment*,

$[a, b[$, $]a, b]$, $[a, +\infty[$ et $]-\infty, b]$ des *intervalles semi-ouverts*,

$]a, b[$, $]a, +\infty[$, $]-\infty, b[$ et $]-\infty, +\infty[$ des *intervalles ouverts*.

Les intervalles $[a, b]$, $]a, b]$, $[a, b[$ et $]a, b[$ sont *finis* et a et b sont leurs *bornes*. Les autres intervalles sont *infinis*.

Aux intervalles numériques sont associés des intervalles sur l'axe de coordonnées. Ainsi l'intervalle $[x_1, x_2]$ est représenté sur l'axe de coordonnées par le segment M_1M_2, où M_1 est le point de coordonnée x_1 et M_2, de coordonnée x_2 (fig. 5). L'ensemble $]-\infty, +\infty[$ est représenté par l'axe de coordonnées tout entier. C'est pourquoi l'ensemble $]-\infty, +\infty[$ s'appelle aussi *axe numérique* et tout nombre, *point* de cet axe. Soient a un point arbitraire de l'axe numérique, $\delta > 0$ un nombre. L'intervalle $]a - \delta, a + \delta[$ s'appelle δ-*voisinage* du point a.

§ 4. Bornes des ensembles numériques

On dit qu'un ensemble X est majoré ou borné supérieurement (resp. minoré ou borné inférieurement) s'il existe un nombre c tel que $x \leqslant c$ (resp. $x \geqslant c$) pour tout $x \in X$. Le nombre c s'appelle *majorant* (resp. *minorant*) de X.

Un ensemble qui est majoré et minoré est dit *borné*.

Par exemple, tout intervalle fini $(]a, b]$, $[a, b[$, $]a, b[$, $]a, b[)$ est borné. L'intervalle $]a, +\infty[$ est minoré mais pas majoré; l'axe numérique $]-\infty, +\infty[$ est un ensemble qui n'est ni minoré ni majoré.

Il est évident que tout ensemble majoré (resp. minoré) admet une infinité de majorants (resp. de minorants). En effet, si un nombre c est un majorant (resp. minorant) de X, tout nombre c' strictement supérieur (resp. strictement inférieur) à c sera aussi un majorant (resp. un minorant) de X, puisque $x \leqslant c$ (resp. $x \geqslant c$) entraîne $x \leqslant c'$ (resp. $x \geqslant c'$).

Il est naturel de se poser la question de l'existence du plus petit majorant d'un ensemble majoré et du plus grand minorant d'un ensemble minoré.

Le plus petit majorant d'un ensemble majoré X s'appelle *borne supérieure* ou *supremum* de X et se note $\sup X$; le plus grand minorant, *borne inférieure* ou *infimum* de X et se note $\inf X$.

Exemples. Soit $X =]a, b[$. Le nombre b est la borne supérieure de X, le nombre a, sa borne inférieure, autrement dit $b = \sup X$, $a = \inf X$. Soit $X =]a, +\infty[$. Dans ce cas $a = \inf X$, quant aux majorants et à la borne supérieure, ils n'existent pas.

La borne supérieure jouit de l'importante propriété suivante: si petit que soit le nombre $\varepsilon > 0$, il existe un $x \in X$ tel que $x > \sup X - \varepsilon$ *). En effet, s'il n'existait pas un tel x, le nombre $\sup x$ ne serait pas borne supérieure (c'est-à-dire le plus petit majorant). En d'autres termes, cette propriété exprime que le nombre $\sup X$ est le plus petit de tous les majorants de X et il ne peut être diminué.

Cette propriété de la borne supérieure peut être reformulée dans les termes suivants: si $c = \sup X$, pour tout nombre $c' < c$ il existe un nombre $x \in X$ tel que $x > c'$ **). Pour s'assurer de l'équivalence de ces deux formulations, il suffit de prendre c' et ε tels que $c' = c - \varepsilon$, relation qui entraîne l'équivalence de la condition $\varepsilon > 0$ à la condition $c' < c$.

La borne inférieure jouit de la même propriété: si petit que soit le nombre $\varepsilon > 0$, on peut exhiber un $x \in X$ tel que $x < \inf X + \varepsilon$. (On demande au lecteur de formuler cette propriété sous une autre forme.)

*) Ou, en signes logiques: $\forall \varepsilon > 0$, $\exists x \in X$: $x > \sup X - \varepsilon$.

**) Ou, en signes logiques: $\forall c' < c$, $\exists x \in X$: $x > c'$.

Il se pose la question de savoir si la borne supérieure (resp. inférieure) d'un ensemble majoré (resp. minoré) existe toujours. La réponse à cette question est donnée par l'important théorème suivant.

Théorème 1.1. *Tout ensemble numérique majoré* (resp. *minoré*) *non vide admet une borne supérieure* (resp. *inférieure*).

Démonstration. Soit X un ensemble majoré non vide. Alors l'ensemble Y des majorants de X n'est pas vide. De la définition de la borne supérieure il s'ensuit que pour tout $x \in X$ et tout $y \in Y$ on a $x \leqslant y$. D'après la propriété de continuité des nombres réels il existe un nombre c tel que, quels que soient x et y,

$$x \leqslant c \leqslant y. \tag{1}$$

De la première inégalité (1) il résulte que c est un majorant de X, de la deuxième, qu'il est le plus petit majorant, c'est-à-dire la borne supérieure.

On démontre de façon analogue l'existence de la borne inférieure d'un ensemble minoré non vide. ■

Si un ensemble X n'est pas majoré (resp. minoré), on convient d'écrire: $\sup X = +\infty$ (resp. $\inf X = -\infty$).

§ 5. Valeur absolue d'un nombre

La notion de valeur absolue et les inégalités relatives aux valeurs absolues interviennent souvent dans la suite.

Définition. On appelle *valeur absolue* (ou *module*) d'un nombre x le nombre x lui-même ou son opposé $-x$ selon que $x \geqslant 0$ ou $x < 0$.

La valeur absolue de x se note $|x|$. Ainsi

$$|x| = \begin{cases} x \text{ si } x \geqslant 0, \\ -x \text{ si } x < 0. \end{cases}$$

Les propriétés suivantes de la valeur absolue résultent directement de sa définition.

1° $|x| \geqslant 0$. En effet:

si $x \geqslant 0$, alors $|x| = x \geqslant 0$;

si $x < 0$, on a $|x| = -x$; or $-x > 0$, puisque $x < 0$, autrement dit $|x| > 0$. Dans les deux cas on a $|x| \geqslant 0$.

2° $|x| = |-x|$. En effet:

si $x \geqslant 0$, alors $-x \leqslant 0$ et $|-x| = -(-x) = x = |x|$;

si $x < 0$, alors $-x > 0$ et $|-x| = -x = |x|$, puisque $x < 0$. Dans les deux cas on a $|x| = |-x|$.

3° $-|x| \leqslant x \leqslant |x|$. En effet:

si $x \geqslant 0$, alors $|x| = x$ et $-x \leqslant 0$. D'où $|x| \geqslant -x$, c'est-à-dire que $-|x| \leqslant x = |x|$;

si $x < 0$, alors $|x| = -x$, d'où $-|x| = x$. Comme $x < 0$, on a $2x < 0$, ou $x + x < 0$, d'où $x < -x$, c'est-à-dire que $x < |x|$. Ainsi $-|x| = x < |x|$. Dans les deux cas on a $-|x| \leqslant x \leqslant |x|$.

Les trois propriétés suivantes seront établies sous forme de théorèmes en raison de leur importance.

Théorème 1.2. *Soit $\varepsilon > 0$ un nombre. Les inégalités $|x| \leqslant \varepsilon$ et $-\varepsilon \leqslant x \leqslant \varepsilon$ sont alors équivalentes* *).

Démonstration. Soit $|x| \leqslant \varepsilon$. Alors:

1) si $x \geqslant 0$, alors $|x| = x$, donc $x \leqslant \varepsilon$, d'où $0 \leqslant x \leqslant \varepsilon$;

2) si $x < 0$, alors $|x| = -x$, donc $-x \leqslant \varepsilon$, d'où $-\varepsilon \leqslant x \leqslant 0$. En combinant 1) et 2) on trouve que $-\varepsilon \leqslant x \leqslant \varepsilon$ pour tout x.

Supposons maintenant que $-\varepsilon \leqslant x \leqslant \varepsilon$. Ceci exprime qu'on a simultanément $x \leqslant \varepsilon$ et $x \geqslant -\varepsilon$. La dernière inégalité nous donne $-x \leqslant \varepsilon$. Puisque par définition $|x|$ est égal soit à x, soit à $-x$, il vient $|x| \leqslant \varepsilon$. ■

Théorème 1.3. *La valeur absolue de la somme de deux nombres est inférieure à la somme des valeurs absolues de ces nombres, c'est-à-dire que $|x + y| \leqslant |x| + |y|$.*

Démonstration. Soient x et y des nombres quelconques. D'après la propriété 3°, on a

$$-|x| \leqslant x \leqslant |x| \quad \text{et} \quad -|y| \leqslant y \leqslant |y|.$$

En ajoutant membre à membre, on obtient

$$-(|x| + |y|) \leqslant x + y \leqslant (|x| + |y|).$$

En vertu du théorème 1.2 cette double inégalité équivaut à l'inégalité $|x + y| \leqslant |x| + |y|$. ■

Remarquons que $|x - y| \leqslant |x| + |y|$.

Théorème 1.4. *La valeur absolue de la différence de deux nombres est supérieure à la différence des valeurs absolues de ces nombres, soit $|x - y| \geqslant |x| - |y|$.*

Démonstration. Quels que soient x et y, on a

$$x = y + (x - y).$$

D'après le théorème 1.3

$$|x| = |y + (x - y)| \leqslant |y| + |x - y|.$$

D'où $|x - y| \geqslant |x| - |y|$. ■

Remarquons que $|x + y| \geqslant |x| - |y|$.

*) En signes logiques cette assertion s'écrit: $\forall \varepsilon > 0: |x| < \varepsilon \Leftrightarrow \Leftrightarrow -\varepsilon < x < \varepsilon$.

Signalons en conclusion que quels que soient les nombres x et y, on a les relations (qui se vérifient sans peine):

$$|x \cdot y| = |x| \cdot |y| \quad \text{et} \quad \left|\frac{x}{y}\right| = \frac{|x|}{|y|} \text{ si } y \neq 0.$$

§ 6. Raisonnement par récurrence

Le raisonnement par récurrence est un outil très puissant en mathématiques qui est utilisé dans la démonstration de propositions faisant intervenir un entier naturel n. La procédure générale est la suivante: pour prouver une proposition mettant en jeu un entier naturel n (par exemple une formule), il faut: 1) vérifier qu'elle est vraie pour $n = 1$ *); 2) supposer qu'elle est vraie pour n et la prouver pour $n + 1$. Ceci fait, on conclut que cette proposition est vraie quel que soit n. Voyons des exemples.

Exemple 1. Prouver par récurrence que

$$1^2 + 2^2 + 3^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

S o l u t i o n . 1) Assurons-nous que cette formule est vraie pour $n = 1$. Le premier membre est égal à 1, le second à $\frac{1(1+1)(2\cdot 1+1)}{6} = 1$. Cette formule est donc vraie pour $n = 1$. 2) Supposons qu'elle est vraie pour n et prouvons-la pour $n+1$, autrement dit, prouvons que

$$1^2 + 2^2 + 3^2 + \ldots + n^2 + (n+1)^2 = \frac{(n+1)[(n+1)+1][2(n+1)+1]}{6}.$$

En effet,

$$1^2 + 2^2 + 3^2 + \ldots + n^2 + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + (n+1)^2 =$$

$$= \frac{(n+1)[n(2n+1)+6(n+1)]}{6} = \frac{(n+1)(2n^2+7n+6)}{6} =$$

$$= \frac{(n+1)(n+2)(2n+3)}{6} = \frac{(n+1)[(n+1)+1][2(n+1)+1]}{6}.$$

C.Q.F.D. Cette formule est donc vraie quel que soit n.

Le raisonnement par récurrence se prête bien au calcul des sommes d'un nombre fini de termes.

Exemple 2. Trouver la somme

$$1 + 3 + 5 + \ldots + (2n - 1).$$

S o l u t i o n. Désignons cette somme par S_n:

$$S_n = 1 + 3 + 5 + \ldots + (2n - 1).$$

*) Si la proposition n'a pas de sens pour $n = 1$, il faut la vérifier pour la plus petite valeur de n pour laquelle elle a un sens.

Calculons les premières valeurs de cette somme:

$$S_1 = 1;\ S_2 = 1 + 3 = 4;\ S_3 = 1 + 3 + 5 = 9;$$
$$S_4 = 1 + 3 + 5 + 7 = 16.$$

Nous remarquons que $S_n = n^2$. Raisonnons par récurrence pour prouver cette assertion. Supposons que cette formule est vraie pour n et prouvons-la pour $n + 1$, c'est-à-dire que $S_{n+1} = (n + 1)^2$. En effet,

$$S_{n+1} = S_n + (2(n + 1) - 1) = n^2 + (2n + 1) = (n + 1)^2,$$

C.Q.F.D. La formule $S_n = n^2$ est donc vraie pour tout n et

$$1 + 3 + 5 + \ldots + (2n - 1) = n^2.$$

Exercice. Calculer la somme $\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \ldots + \frac{1}{n(n+1)}$. $\left(\text{Réponse: } \frac{n}{n+1}\right)$.

CHAPITRE II

LIMITE D'UNE SUITE

Les notions de limite et de fonction sont fondamentales en analyse mathématique. La notion de limite est étudiée en mathématiques élémentaires où elle permet le calcul de la longueur de la circonférence, du volume du cylindre, du cône, etc. Elle est également utilisée pour calculer la somme d'une progression géométrique infiniment décroissante. Le passage à la limite est une opération fondamentale en analyse. Dans ce chapitre on envisage une forme élémentaire de passage à la limite, basée sur la notion de limite d'une suite numérique. Cette notion nous permettra de définir ultérieurement des formes plus complexes de passage à la limite.

§ 1. Suites numériques

1. Suites numériques et opérations arithmétiques sur elles. Les suites numériques sont étudiées dans le secondaire. Comme exemples de suites citons : 1) la suite des termes d'une progression arithmétique ou géométrique ; 2) la suite des périmètres de polygones réguliers inscrits dans un cercle ; 3) la suite $x_1 = 1$, $x_2 = 1{,}4$, $x_3 = 1{,}41\ldots$ des valeurs approchées de $\sqrt{2}$.

Précisons et élargissons la notion de suite numérique.

Définition. Si à chaque nombre n de la suite naturelle

$$1, 2, 3, \ldots, n, \ldots$$

on associe un nombre réel x_n, on obtient un ensemble de réels

$$x_1, x_2, x_3, \ldots, x_n, \ldots \qquad (1)$$

appelé *suite numérique* ou simplement *suite* *).

Les nombres x_1, x_2, x_3, $\ldots$, x_n, $\ldots$ s'appellent *éléments* (ou *termes*) de la suite (1), x_n, l'*élément* (ou *terme*) *général* de la sui-

*) En d'autres termes, on peut définir une suite numérique comme un ensemble de couples $(n ; x_n)$ dont le premier nombre prend successivement les valeurs 1, 2, 3, ...

te, n, l'*indice* ou le *rang* de x_n. La suite (1) sera notée $\{x_n\}$. Ainsi $\left\{\frac{1}{n}\right\}$ est la suite $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$

On considère qu'une suite est donnée si l'on sait calculer n'importe lequel de ses éléments. Par exemple, la formule $x_n = 1 + (-1)^n$ définit la suite: 0, 2, 0, 2, . . . En convertissant la fraction $\frac{1}{3}$ en nombre décimal et en gardant une, deux, trois, etc., décimales, on obtient la suite

$$x_1 = 0{,}3;\ x_2 = 0{,}33;\ x_3 = 0{,}333;\ \ldots;\ x_n = 0{,}333\ldots3;\ \ldots$$

Toute suite contient par définition une infinité d'éléments: deux quelconques de ses éléments se différencient au moins par leur rang.

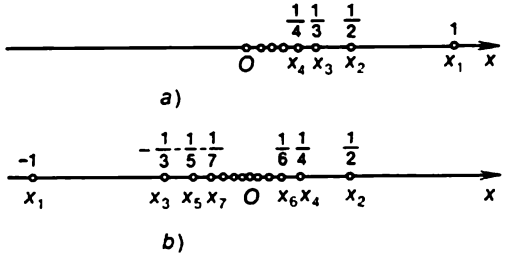


Fig. 6

Géométriquement une suite se représente sur l'axe de coordonnées par une infinité de points dont les coordonnées sont égales aux éléments respectifs de cette suite. Les figures 6, *a* et *b* représentent respectivement les suites $\{x_n\} = \left\{\frac{1}{n}\right\}$ et $\{x_n\} = \left\{\frac{(-1)^n}{n}\right\}$.

Introduisons les opérations arithmétiques sur les suites numériques. Soient données les suites $\{x_n\}$ et $\{y_n\}$.

On appelle:

produit de la suite $\{x_n\}$ *par un nombre* m la suite $\{mx_n\}$;

somme des suites $\{x_n\}$ et $\{y_n\}$ la suite $\{x_n + y_n\}$;

différence des suites $\{x_n\}$ et $\{y_n\}$ la suite $\{x_n - y_n\}$;

produit des suites $\{x_n\}$ et $\{y_n\}$ la suite $\{x_n y_n\}$;

quotient des suites $\{x_n\}$ et $\{y_n\}$ la suite $\left\{\frac{x_n}{y_n}\right\}$, pourvu que tous les termes de la suite $\{y_n\}$ soient non nuls.

2. Suites bornées et illimitées.

Définition 1. On dit qu'une suite $\{x_n\}$ est *majorée* (resp. *minorée*) s'il existe un nombre M (resp. m) tel que pour tout élément x_n on a $x_n \leqslant M$ (resp. $x_n \geqslant m$).

Définition 2. On dit qu'une suite $\{x_n\}$ est *bornée* si elle est à la fois majorée et minorée, c'est-à-dire s'il existe des nombres M et m tels que pour tout élément x_n on a $m \leqslant x_n \leqslant M$.

Soit $A = \max \{ |m|, |M| \}$. La condition que la suite $\{x_n\}$ est bornée peut alors s'écrire $|x_n| \leqslant A$.

Définition 3. On dit qu'une suite $\{x_n\}$ est *illimitée* si aussi grand que soit le nombre $A > 0$ il existe un élément x_n tel que $|x_n| > A$ (autrement dit, soit $x_n > A$, soit $x_n < -A$).

De ces définitions il s'ensuit que si une suite est majorée, tous ses éléments appartiennent à l'intervalle $]-\infty, M]$; si elle est minorée, à l'intervalle $[m, +\infty[$; si elle est bornée, à l'intervalle $[m, M]$. Une suite illimitée peut être majorée (ou minorée).

Considérons des exemples de suites bornées et illimitées.

1. La suite $1, 2, 3, \ldots, n, \ldots$ est minorée mais pas majorée.

2. La suite $-1, -2, -3, \ldots, -n, \ldots$ est majorée mais pas minorée.

3. La suite $1, 1/2, 1/3, \ldots, 1/n, \ldots$ est bornée, puisque tout élément x_n est tel que $0 \leqslant x_n \leqslant 1$ $(m = 0, M = 1)$.

4. La suite $-1, 2, -3, 4, -5, \ldots, (-1)^n n, \ldots$ est illimitée. En effet, si grand que soit le nombre A, il existe des éléments x_n de cette suite tels que $|x_n| > A$.

Ces définitions s'écrivent de la manière suivante à l'aide des signes logiques:

la suite $\{x_n\}$ est majorée si $\exists M, \forall x_n : x_n \leqslant M$;

minorée si $\exists m, \forall x_n : x_n \geqslant m$;

bornée si $\exists A > 0, \forall x_n : |x_n| \leqslant A$;

illimitée si $\forall A > 0, \exists x_n : |x_n| > A$.

Si l'on compare les notations logiques des deux dernières définitions, on constate que les quantificateurs $\exists$ et $\forall$ se substituent l'un à l'autre dans les négations.

3. Suites infiniment grandes et infiniment petites.

Définition 1. On dit qu'une suite $\{x_n\}$ est *infiniment grande* si pour tout nombre $A > 0$ si grand soit-il il existe un entier N tel que pour $n > N$ on ait $|x_n| > A$.

Cette définition s'écrit en signes logiques:

$$\forall A > 0, \exists N, \forall n > N : |x_n| > A.$$

Remarque. Il est évident que toute suite infiniment grande est illimitée. Mais une suite illimitée n'est pas nécessairement infiniment grande. Ainsi la suite illimitée $1, 2, 1, 3, \ldots, 1, n, 1, n+1, \ldots$ n'est pas infiniment grande, puisque pour $A > 1$ l'inégalité $|x_n| > A$ n'est pas réalisée pour les termes x_n de rang impair.

Définition 2. On dit qu'une suite $\{\alpha_n\}$ est *infiniment petite* si pour tout $\varepsilon > 0$ il existe un entier N tel que pour $n > N$ on a $|\alpha_n| < \varepsilon$.

En signes logiques cette définition s'ecrit:

$$\forall \varepsilon > 0,\ \exists N,\ \forall n > N:\ |\alpha_n| < \varepsilon.$$

Exemple 1. Utilisons la définition 1 pour prouver que la suite $\{n\}$ est infiniment grande.

Prenons un nombre $A > 0$. De l'inégalité $|x_n| = |n| > A$ on déduit que $n > A$. Si l'on prend $N \geqslant A$, on aura $|x_n| > A$ pour tout $n > N$, ce qui exprime en vertu de la définition 1 que la suite $\{n\}$ est infiniment grande.

Exemple 2. Utilisons la définition 2 pour prouver que la suite $\{1/n\}$ est infiniment petite.

Soit $\varepsilon > 0$ un nombre quelconque. L'inégalité $|\alpha_n| = |1/n| < \varepsilon$ entraîne que $n > 1/\varepsilon$. Si l'on prend $N = [1/\varepsilon]$ *), pour tout $n > N$ on aura $n \geqslant [1/\varepsilon] + 1 > 1/\varepsilon$, d'où $1/n = |\alpha_n| < \varepsilon$. La suite $\{1/n\}$ est donc infiniment petite d'après la définition 2.

Prouvons un théorème établissant un lien entre les suites infiniment grandes et les suites infiniment petites.

Théorème 2.1. *Si $\{x_n\}$ est une suite infiniment grande dont les termes sont tous différents de zéro, la suite $\left\{\frac{1}{x_n}\right\}$ est infiniment petite, et réciproquement, si $\{\alpha_n\}$ est une suite infiniment petite et $\alpha_n \neq 0$, la suite $\left\{\frac{1}{\alpha_n}\right\}$ est infiniment grande.*

Démonstration. Soit $\{x_n\}$ une suite infiniment grande. Considérons un $\varepsilon > 0$ quelconque et posons $A = \frac{1}{\varepsilon}$. D'après la définition 1 pour cet A il existe un entier N tel que pour $n > N$ on a $|x_n| > A$. D'où $\left|\frac{1}{x_n}\right| = \frac{1}{|x_n|} < \frac{1}{A} = \varepsilon$ pour tout $n > N$. Or ceci exprime que la suite $\left\{\frac{1}{x_n}\right\}$ est infiniment petite.

La deuxième partie du théorème se démontre de façon analogue. ■

4. Propriétés fondamentales des suites infiniment petites.

Théorème 2.2. *La somme et la différence de deux suites infiniment petites sont des suites infiniment petites.*

Démonstration. Soient $\{\alpha_n\}$ et $\{\beta_n\}$ des suites infiniment petites. On demande de prouver que les suites $\{\alpha_n \pm \beta_n\}$ le sont aussi. Soient $\varepsilon > 0$ un nombre arbitraire, N_1 un entier tel

*) Le symbole $[x]$ désigne la partie entière de x, c'est-à-dire le plus grand entier inférieur à x. Par exemple, $[1] = 1$, $[3{,}1] = 3$, $[0{,}7] = 0$, $[-0{,}5] = -1$, $[-172{,}9] = -173$, etc. Il est évident que $[x] + 1 > x$.

que $|\alpha_n| < \frac{\varepsilon}{2}$ pour $n > N_1$, N_2, un entier tel que $|\beta_n| < \frac{\varepsilon}{2}$ pour $n > N_2$. (Les entiers N_1 et N_2 existent par définition d'une suite infiniment petite.) Prenons $N = \max\{N_1, N_2\}$; pour $n > N$ on aura alors simultanément: $|\alpha_n| < \frac{\varepsilon}{2}$, $|\beta_n| < \frac{\varepsilon}{2}$. Donc pour $n > N$ on a

$$|\alpha_n \pm \beta_n| \leqslant |\alpha_n| + |\beta_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Ce qui exprime que les suites $\{\alpha_n \pm \beta_n\}$ sont infiniment petites. ■

Corollaire. *La somme algébrique d'un nombre fini de suites infiniment petites est une suite infiniment petite.*

Théorème 2.3. *Le produit de deux suites infiniment petites est une suite infiniment petite.*

Démonstration. Soient $\{\alpha_n\}$ et $\{\beta_n\}$ des suites infiniment petites. On demande de prouver qu'il en est de même de la suite $\{\alpha_n \cdot \beta_n\}$. La suite $\{\alpha_n\}$ étant infiniment petite, pour tout $\varepsilon > 0$ il existe un entier N_1 tel que $|\alpha_n| < \varepsilon$ pour $n > N_1$; la suite $\{\beta_n\}$ étant aussi infiniment petite, il existe pour $\varepsilon = 1$ un entier N_2 tel que $|\beta_n| < 1$ pour $n > N_2$. Prenons $N = \max\{N_1, N_2\}$; ces deux inégalités seront alors réalisées pour $n > N$. Donc pour $n > N$

$$|\alpha_n \cdot \beta_n| = |\alpha_n| \cdot |\beta_n| < \varepsilon \cdot 1 = \varepsilon.$$

Ce qui exprime que la suite $\{\alpha_n \cdot \beta_n\}$ est infiniment petite. ■

Corollaire. *Le produit d'un nombre fini de suites infiniment petites est une suite infiniment petite.*

Remarque. Le quotient de deux suites infiniment petites peut ne pas être une suite infiniment petite et peut même ne pas avoir de sens. Par exemple, si $\alpha_n = 1/n$ et $\beta_n = 1/n$, les éléments de $\{\alpha_n/\beta_n\}$ sont tous égaux à 1 et cette suite est bornée. Si $\alpha_n = 1/n$ et $\beta_n = 1/n^2$, la suite $\{\alpha_n/\beta_n\}$ est infiniment grande, la suite $\{\beta_n/\alpha_n\}$, infiniment petite. La suite $\{\alpha_n/\beta_n\}$ n'a pas de sens si à partir d'un certain rang tous les éléments de $\{\beta_n\}$ sont nuls.

Théorème 2.4. *Le produit d'une suite bornée par une suite infiniment petite est une suite infiniment petite.*

Démonstration. Soient $\{x_n\}$ une suite bornée, $\{\alpha_n\}$, une suite infiniment petite. On demande de prouver que la suite $\{x_n \cdot \alpha_n\}$ est infiniment petite. La suite $\{x_n\}$ étant bornée, il existe un $A > 0$ tel que tout élément x_n satisfait l'inégalité $|x_n| \leqslant A$. Soit $\varepsilon > 0$ un nombre quelconque. La suite $\{\alpha_n\}$ étant infiniment petite, pour tout nombre ε/A il existe un entier

N tel que $|\alpha_n| < \varepsilon/A$ pour $n > N$. Donc, pour $n > N$

$$|x_n \cdot \alpha_n| = |x_n| \cdot |\alpha_n| < A \cdot \frac{\varepsilon}{A} = \varepsilon.$$

Ce qui exprime que la suite $\{x_n \cdot \alpha_n\}$ est infiniment petite. ■

Corollaire. *Le produit d'une suite infiniment petite par un nombre est une suite infiniment petite.*

Passons maintenant à l'une des plus importantes notions d'analyse mathématique, la notion de limite d'une suite numérique.

§ 2. Suites convergentes

1. Notion de suite convergente.

Définition. On dit qu'un nombre a est la *limite d'une suite* $\{x_n\}$ si pour tout nombre $\varepsilon > 0$ il existe un entier N tel que pour $n > N$ on ait

$$|x_n - a| < \varepsilon. \tag{1}$$

Cette définition s'écrit à l'aide des signes logiques :

$$\forall \varepsilon > 0, \quad \exists N, \quad \forall n > N : |x_n - a| < \varepsilon.$$

Une suite qui possède une limite s'appelle *convergente*.
Si une suite $\{x_n\}$ converge vers un nombre a, on note ceci :

$$\lim_{n \to \infty} x_n = a \text{ ou } x_n \to a \text{ lorsque } n \to \infty. \tag{2}$$

Une suite qui n'est pas convergente est dite *divergente*.

Exemple. Utilisons la définition de la limite d'une suite pour prouver que $\lim\limits_{n \to \infty} \frac{n}{n+1} = 1$.

Soit $\varepsilon > 0$ un nombre arbitraire. Puisque $|x_n - 1| = \left|\frac{n}{n+1} - 1\right| = \frac{1}{n+1}$, pour trouver les valeurs de n vérifiant l'inégalité $|x_n - 1| < \varepsilon$, il suffit de résoudre l'inéquation $1/(n+1) < \varepsilon$, d'où il résulte que $n > (1 - \varepsilon)/\varepsilon$. Donc pour N on peut prendre la partie entière de $(1 - \varepsilon)/\varepsilon$, soit $N = [(1 - \varepsilon)/\varepsilon]$. L'inégalité $|x_n - 1| < \varepsilon$ sera satisfaite pour tout $n > N$. Ce qui prouve que $\lim\limits_{n \to \infty} \frac{n}{n+1} = 1$.

Remarque 1. Supposons qu'une suite $\{x_n\}$ admet un nombre a pour limite. Alors $\{\alpha_n\} = \{x_n - a\}$ est une suite infiniment petite, puisque pour tout $\varepsilon > 0$ il existe un entier N tel que pour tout $n > N$ on a $|\alpha_n| = |x_n - a| < \varepsilon$. Donc tout terme x_n d'une suite convergeant vers un nombre a peut être représenté sous la forme

$$x_n = a + \alpha_n, \tag{3}$$

où α_n est le terme général de la suite infiniment petite $\{\alpha_n\}$. Il est évident que la réciproque est vraie : si x_n peut être représenté sous la forme $x_n = a + \alpha_n$, où $\{\alpha_n\}$ est une suite infiniment petite, alors $\lim\limits_{n\to\infty} x_n = a$. La représentation (3) est utilisée dans la démonstration des théorèmes sur les limites de suites.

Remarque 2. L'inégalité (1) équivaut aux doubles inégalités

$$-\varepsilon < x_n - a < \varepsilon \text{ ou } a - \varepsilon < x_n < a + \varepsilon,$$

qui expriment que le terme x_n est situé dans un ε-voisinage du point a (fig. 7). On peut donc formuler la définition de la limite d'une suite de la manière suivante : on dit qu'un nombre a est la limite d'une suite $\{x_n\}$ si pour tout ε-voisinage de a il existe un entier N tel que tous les termes x_n de rang $n > N$ se trouvent dans cet ε-voisinage.

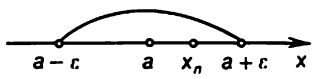


Fig. 7

Remarque 3. Il est évident qu'une suite $\{x_n\}$ infiniment grande ne possède pas de limite. On dit quelquefois qu'elle possède une limite infinie et on écrit

$$\lim_{n\to\infty} x_n = \infty.$$

Si à partir d'un certain rang tous les termes de cette suite sont strictement positifs (resp. strictement négatifs) on note

$$\lim_{n\to\infty} x_n = +\infty \text{ (resp. } \lim_{n\to\infty} x_n = -\infty).$$

La limite d'une suite telle qu'elle a été définie plus haut sera parfois appelée *limite finie* par opposition à la *limite infinie*.

Remarque 4. Il est évident que toute suite infiniment petite est convergente et a pour limite le nombre $a = 0$.

2. Propriétés fondamentales des suites convergentes. Prouvons un lemme qui nous sera utile dans la démonstration du théorème 2.5.

Lemme 2.1. *Si tous les termes d'une suite infiniment petite* $\{\alpha_n\}$ *sont égaux à un même nombre* c, *alors* $c = 0$.

Démonstration. Supposons par absurde que $c \neq 0$. Posons $\varepsilon = \frac{|c|}{2}$. Par définition d'une suite infiniment petite, il existe un entier N tel que pour $n > N$ on a $|\alpha_n| < \varepsilon$. Puisque $\alpha_n = c$ et $\varepsilon = \frac{|c|}{2}$, la dernière inégalité peut être mise sous la

forme $|c| < \frac{|c|}{2}$, d'où l'on déduit que $1 < 1/2$. Cette contradiction prouve que $c = 0$. ■

Théorème 2.5. *Une suite convergente n'a qu'une seule limite.*

Démonstration. Supposons par absurde qu'une suite convergente $\{x_n\}$ admet deux limites distinctes, a et b. D'après la formule (3) on a alors

$$x_n = a + \alpha_n \text{ et } x_n = b + \beta_n,$$

où α_n et β_n sont les termes généraux des suites infiniment petites $\{\alpha_n\}$ et $\{\beta_n\}$. En égalant les seconds membres de ces égalités on trouve que $\alpha_n - \beta_n = b - a$. Puisque tous les termes de la suite infiniment petite $\{\alpha_n - \beta_n\}$ sont égaux à $b - a$, on a $b - a = 0$ en vertu du lemme 2.1. Donc $b = a$. ■

Théorème 2.6. *Une suite convergente est bornée.*

Démonstration. Soit $\{x_n\}$ une suite convergeant vers un nombre a. Soient par ailleurs $\varepsilon > 0$ un nombre arbitraire et N un entier tel que pour $n > N$ on a $|x_n - a| < \varepsilon$. Alors

$$|x_n| = |(x_n - a) + a| \leqslant |x_n - a| + |a| < |a| + \varepsilon$$

pour tout $n > N$. Soit $A = \max \{|a| + \varepsilon, |x_1|, |x_2|, \ldots, |x_N|\}$. Il est évident que $|x_n| \leqslant A$ pour tout n, ce qui exprime que la suite $\{x_n\}$ est bornée. ■

Remarque. Une suite bornée peut ne pas converger. Par exemple la suite $-1, 1, -1, \ldots, (-1)^n, \ldots$ est visiblement bornée, mais pas convergente. Prouvons-le. Supposons que cette suite tend vers un nombre a. Pour $\varepsilon = 1/2$ il existerait alors un entier N tel que pour $n > N$ on aurait $|x_n - a| < 1/2$. Puisque x_n prend alternativement les valeurs 1 et -1, on aurait $|1 - a| < 1/2$ et $|(-1) - a| < 1/2$. En se servant de ces inégalités on obtiendrait

$$2 = |1 - a + a - (-1)| \leqslant |1 - a| + |a - (-1)| < 1/2 + 1/2 = 1,$$

c'est-à-dire que $2 < 1$. Cette contradiction prouve la divergence de la suite $\{(-1)^n\}$.

Théorème 2.7. *La somme* (resp. *la différence*) *de deux suites convergentes* $\{x_n\}$ *et* $\{y_n\}$ *est une suite convergente dont la limite est égale à la somme* (resp. *la différence*) *des limites de* $\{x_n\}$ *et* $\{y_n\}$.

Démonstration. Soient a et b les limites respectives des suites $\{x_n\}$ et $\{y_n\}$. D'après la formule (3) on a alors

$$x_n = a + \alpha_n, \; y_n = b + \beta_n,$$

où $\{\alpha_n\}$ et $\{\beta_n\}$ sont des suites infiniment petites. Donc

$$(x_n \pm y_n) - (a \pm b) = \alpha_n \pm \beta_n.$$

La suite $\{\alpha_n \pm \beta_n\}$ est infiniment petite d'après le théorème 2.2. Donc il en est de même de la suite $\{(x_n \pm y_n) - (a \pm b)\}$ et la suite $\{x_n \pm y_n\}$ converge vers $a \pm b$. ■

Théorème 2.8. *Le produit de deux suites convergentes* $\{x_n\}$ *et* $\{y_n\}$ *est une suite convergente dont la limite est égale au produit des limites de* $\{x_n\}$ et $\{y_n\}$.

Démonstration. Soient a et b les limites respectives des suites $\{x_n\}$ *et* $\{y_n\}$. D'après la formule (3)

$$x_n = a + \alpha_n, \quad y_n = b + \beta_n,$$

où $\{\alpha_n\}$ et $\{\beta_n\}$ sont des suites infiniment petites. Donc

$$x_n y_n - ab = a\beta_n + b\alpha_n + \alpha_n\beta_n.$$

La suite $\{a\beta_n + b\alpha_n + \alpha_n\beta_n\}$ est infiniment petite en vertu des théorèmes 2.2 à 2.4. Donc il en est de même de la suite $\{x_n y_n - ab\}$ et la suite $\{x_n y_n\}$ converge vers le nombre ab. ■

Théorème 2.9. *Le quotient de deux suites convergentes* $\{x_n\}$ *et* $\{y_n\}$ *sous réserve que la limite de* $\{y_n\}$ *soit différente de* 0 *) *est une suite convergente dont la limite est égale au quotient des limites des suites* $\{x_n\}$ *et* $\{y_n\}$.

Démonstration. Soient a et b $(b \neq 0)$ les limites respectives des suites $\{x_n\}$ et $\{y_n\}$. D'après la formule (3),

$$x_n = a + \alpha_n, \quad y_n = b + \beta_n,$$

où $\{\alpha_n\}$ et $\{\beta_n\}$ sont des suites infiniment petites. Donc

$$\frac{x_n}{y_n} - \frac{a}{b} = \frac{bx_n - ay_n}{by_n} = \frac{b(a+\alpha_n) - a(b+\beta_n)}{by_n} = \frac{1}{y_n}\left(\alpha_n - \frac{a}{b}\beta_n\right).$$

La suite $\left\{\alpha_n - \frac{a}{b}\beta_n\right\}$ est infiniment petite d'après les propriétés des suites infiniment petites. Montrons que $\left\{\frac{1}{y_n}\right\}$ est une suite bornée. Puisque $y_n \to b$ lorsque $n \to \infty$, pour $\varepsilon = \frac{|b|}{2}$ il existe un entier N tel que $n > N$ on aura $|y_n - b| < |b|/2$. Donc

$$|y_n| = |b - (b - y_n)| \geqslant |b| - |y_n - b| > |b| - \frac{|b|}{2} = \frac{|b|}{2},$$

*) D'après la condition $\lim\limits_{n\to\infty} y_n \neq 0$, les termes y_n seront non nuls à partir d'un certain rang N, donc le quotient $\{x_n/y_n\}$ a un sens pour $n > N$.

c'est-à-dire que $|y_n| > |b|/2$ et par suite $\left|\frac{1}{y_n}\right| < \frac{2}{|b|}$ pour tout $n > N$. Ceci exprime que la suite $\left\{\frac{1}{y_n}\right\}$ est bornée.

La suite $\left\{\frac{1}{y_n}\left(\alpha_n - \frac{a}{b}\beta_n\right)\right\}$ est infiniment petite d'après le théorème 2.4. Donc il en est de même de la suite $\left\{\frac{x_n}{y_n} - \frac{a}{b}\right\}$ et la suite $\left\{\frac{x_n}{y_n}\right\}$ converge vers $\frac{a}{b}$. ■

Les théorèmes prouvés dans ce numéro revêtent une grande signification non seulement théorique mais aussi pratique.

Mais malgré leur simplicité ils constituent souvent une pierre d'achoppement pour de nombreux débutants. Une attention toute particulière doit être portée au fait que l'application de ces théorèmes implique l'existence de limites finies. L'oubli de ce fait nous conduit aux erreurs suivantes.

Soit la suite $\left\{\frac{5n+1}{n}\right\}$. D'une part, $\lim\limits_{n\to\infty} \frac{5n+1}{n} = \lim\limits_{n\to\infty}\left(5 + \frac{1}{n}\right) =$ $= \lim\limits_{n\to\infty} 5 + \lim\limits_{n\to\infty} \frac{1}{n} = 5 + 0 = 5$, de l'autre

$$\lim_{n\to\infty} \frac{5n+1}{n} = \frac{\lim\limits_{n\to\infty}(5n+1)}{\lim\limits_{n\to\infty} n} = \frac{\infty}{\infty} = 1.$$

D'où $5 = 1$!?

Soit la suite $\{1\}$. D'une part, $\lim\limits_{n\to\infty} 1 = 1$, de l'autre, $1 =$ $= (n+1) - n$ et

$$\lim_{n\to\infty} 1 = \lim_{n\to\infty}(n+1) - n = \lim_{n\to\infty}(n+1) - \lim_{n\to\infty} n = \infty - \infty = 0.$$

D'où $1 = 0$!?

Soit enfin la suite $\left\{\frac{n-1}{n}\right\}$. D'une part, $\lim\limits_{n\to\infty} \frac{n-1}{n} =$ $= \lim\limits_{n\to\infty}\left(1 - \frac{1}{n}\right) = \lim\limits_{n\to\infty} 1 - \lim\limits_{n\to\infty} \frac{1}{n} = 1 - 0 = 1$, de l'autre,

$$\lim_{n\to\infty} \frac{n-1}{n} = \lim_{n\to\infty}(n-1) \cdot \lim_{n\to\infty} \frac{1}{n} = \infty \cdot 0 = 0.$$

D'où $0 = 1$!?

La même erreur a été commise dans les trois cas: les formules

$$\lim_{n\to\infty}(x_n - y_n) = \lim_{n\to\infty} x_n - \lim_{n\to\infty} y_n,$$

$$\lim_{n\to\infty} \frac{x_n}{y_n} = \frac{\lim\limits_{n\to\infty} x_n}{\lim\limits_{n\to\infty} y_n}; \quad \lim_{n\to\infty} x_n \cdot y_n = \lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} y_n$$

ont été appliquées à des suites $\{x_n\}$ ou $\{y_n\}$ ne possédant pas de limite finie. L'écriture $\lim\limits_{n\to\infty} x_n = \infty$ traduit le fait que les termes de la suite $\{x_n\}$ croissent indéfiniment en valeur absolue. Donc on ne peut manipuler le symbole ∞ comme un nombre ordinaire. En particulier, on ne peut écrire

$$\infty - \infty = 0 \quad \text{ou} \quad \frac{\infty}{\infty} = 1 \quad \text{ou encore} \quad \infty \cdot 0 = 0.$$

Passons maintenant à des exemples plus caractéristiques.

Exemple 1. Calculer $\lim\limits_{n\to\infty} \dfrac{2n^2+n+1}{3n^2-1}$.

Solution. Le théorème de la limite d'un quotient ne passe pas directement ici, puisque le numérateur et le dénominateur ont des limites infinies. Mais si l'on transforme cette suite en divisant le numérateur et le dénominateur par n^2, on peut appliquer les théorèmes de la limite d'un quotient et d'une somme. On trouve

$$\lim_{n\to\infty} \frac{2n^2+n+1}{3n^2-1} = \lim_{n\to\infty} \frac{2+1/n+1/n^2}{3-1/n^2} = \frac{\lim\limits_{n\to\infty}(2+1/n+1/n^2)}{\lim\limits_{n\to\infty}(3-1/n^2)} =$$

$$= \frac{\lim\limits_{n\to\infty} 2 + \lim\limits_{n\to\infty}(1/n) + \lim\limits_{n\to\infty}(1/n^2)}{\lim\limits_{n\to\infty} 3 - \lim\limits_{n\to\infty}(1/n^2)} = \frac{2+0+0}{3-0} = \frac{2}{3}.$$

Exemple 2. Calculer $\lim\limits_{n\to\infty}\left(\dfrac{5n}{n+1} + \dfrac{\sin n}{n}\right)$.

Solution. Considérons d'abord la suite $\left\{\dfrac{5n}{n+1}\right\}$. Comme dans l'exemple précédent, transformons cette suite en divisant le numérateur et le dénominateur par n et appliquons ensuite le théorème de la limite d'un quotient et d'une somme. Nous obtenons:

$$\lim_{n\to\infty} \frac{5n}{n+1} = \lim_{n\to\infty} \frac{5}{1+\frac{1}{n}} = \frac{\lim\limits_{n\to\infty} 5}{\lim\limits_{n\to\infty} 1 + \lim\limits_{n\to\infty} \frac{1}{n}} = \frac{5}{1+0} = 5.$$

La suite $\left\{\dfrac{\sin n}{n}\right\}$ peut être traitée comme le produit de la suite bornée $\{\sin n\}$ (puisque $|\sin n| \leqslant 1$) par la suite infiniment petite $\left\{\dfrac{1}{n}\right\}$. D'après le théorème 2.4 c'est une suite infiniment petite qui a zéro pour limite. En définitive,

$$\lim_{n\to\infty}\left(\frac{5n}{n+1} + \frac{\sin n}{n}\right) = \lim_{n\to\infty} \frac{5n}{n+1} + \lim_{n\to\infty}\left(\sin n \cdot \frac{1}{n}\right) = 5 + 0 = 5.$$

Exemple 3. Calculer $\lim\limits_{n\to\infty} \dfrac{2n-3}{n^2+1}$.

Solution. On a

$$\lim_{n\to\infty}\frac{2n-3}{n^2+1}=\lim_{n\to 0}\frac{2-\frac{3}{n}}{n+\frac{1}{n}}=0,$$

puisque pour $n\to\infty$ la suite $\left\{2-\frac{3}{n}\right\}$ est bornée (prouvez-le!), la suite $\left\{n+\frac{1}{n}\right\}$ est infiniment grande (prouvez-le!) et d'après le théorème 2.1 la suite $\left\{\frac{1}{n+\frac{1}{n}}\right\}$ est infiniment petite. Donc en vertu du théorème 2.4

$$\lim_{n\to\infty}\left[\left(2-\frac{3}{n}\right)\cdot\frac{1}{n+\frac{1}{n}}\right]=0.$$

Exemple 4. Calculer $\lim\limits_{n\to\infty}\frac{2n^3+4}{n^2+5}$.

Solution. On a

$$\lim_{n\to\infty}\frac{2n^3+4}{n^2+5}=\lim_{n\to\infty}\frac{2+\frac{4}{n^3}}{\frac{1}{n}+\frac{5}{n^3}}=\infty;$$

en effet, puisque pour $n\to\infty$ la suite $\{z_n\}=\left\{2+\frac{4}{n^3}\right\}$ est bornée ($z_n\to 2$), la suite $\{y_n\}=\left\{\frac{1}{n}+\frac{5}{n^3}\right\}$ est infiniment petite (prouvez-le!) et $\lim\limits_{n\to\infty}\frac{y_n}{z_n}=\lim\limits_{n\to\infty}y_n\cdot\frac{1}{z_n}=0\left(\frac{1}{z_n}\to\frac{1}{2}\right)$ d'après le théorème 2.4, la suite $\left\{\frac{2n^3+4}{n^2+5}\right\}$ est infiniment grande en vertu du théorème 2.1 et sa limite est égale à ∞.

Exemple 5. Calculer $\lim\limits_{n\to\infty}\frac{1+2+3+\ldots+n}{n^2}$.

Solution. Puisque le numérateur est la somme des termes d'une progression arithmétique de raison $d=1$, somme qui est égale à $\frac{(1+n)n}{2}$, on a

$$\lim_{n\to\infty}\frac{1+2+3+\ldots+n}{n^2}=\lim_{n\to\infty}\frac{(1+n)n}{2n^2}=$$

$$=\lim_{n\to\infty}\frac{n+n^2}{2n^2}=\lim_{n\to\infty}\frac{1+\frac{1}{n}}{2}=\frac{1+0}{2}=\frac{1}{2}.$$

Exemple 6. Calculer $\lim\limits_{n\to\infty}\dfrac{1+\frac{1}{2}+\frac{1}{2^2}+\ldots+\frac{1}{2^n}}{1+\frac{1}{3}+\frac{1}{3^2}+\ldots+\frac{1}{3^n}}$.

Solution. Au numérateur et au dénominateur on reconnaît des progressions géométriques de raison $q = 1/2$ et $q = 1/3$ respectivement. La somme des termes de ces progressions étant égale à $\frac{1-q^{(n+1)}}{1-q}$, on a

$$\lim_{n\to\infty}\frac{1+\frac{1}{2}+\frac{1}{2^2}+\ldots+\frac{1}{2^n}}{1+\frac{1}{3}+\frac{1}{3^2}+\ldots+\frac{1}{3^n}}=\lim_{n\to\infty}\frac{\left(1-\frac{1}{2^{n+1}}\right)\left(1-\frac{1}{3}\right)}{\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3^{n+1}}\right)}=$$

$$=\frac{(1-0)\left(1-\frac{1}{3}\right)}{\left(1-\frac{1}{2}\right)(1-0)}=\frac{\frac{2}{3}}{\frac{1}{2}}=\frac{4}{3}.$$

Exemple 7. Calculer $\lim\limits_{n\to\infty}\left[\frac{1}{1\cdot2}+\frac{1}{2\cdot3}+\ldots+\frac{1}{n(n+1)}\right]$.

Solution. On a

$$\lim_{n\to\infty}\left[\frac{1}{1\cdot2}+\frac{1}{2\cdot3}+\ldots+\frac{1}{n(n+1)}\right]=$$

$$=\lim_{n\to\infty}\left[\frac{2-1}{1\cdot2}+\frac{3-2}{2\cdot3}+\ldots+\frac{n+1-n}{n(n+1)}\right]=$$

$$=\lim_{n\to\infty}\left[1-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\frac{1}{3}-\frac{1}{4}+\ldots+\frac{1}{n-1}-\frac{1}{n}+\right.$$

$$\left.+\frac{1}{n}-\frac{1}{n+1}\right]=\lim_{n\to\infty}\left[1-\frac{1}{n+1}\right]=$$

$$=\lim_{n\to\infty}\frac{n}{n+1}=\lim_{n\to\infty}\frac{1}{1+\frac{1}{n}}=\frac{1}{1+0}=1.$$

Exemple 8. Calculer $\lim\limits_{n\to\infty}\left(\frac{n^3}{n^2+1}-\frac{3n^2}{3n+1}\right)$.

Solution. Le théorème de la limite d'une différence ne peut être appliqué, puisque $\lim\limits_{n\to\infty}\frac{n^3}{n^2+1}=\infty$ (prouvez-le !) et $\lim\limits_{n\to\infty}\frac{3n^2}{3n+1}=\infty$ (prouvez-le !). Transformons donc l'expression figurant sous le signe lim en la ramenant au même dénominateur et en simplifiant par n^3. En appliquant ensuite les théorèmes de la limite d'un quotient, d'un produit et d'une différence, on

trouve :

$$\lim_{n\to\infty}\left(\frac{n^3}{n^2+1}-\frac{3n^2}{3n+1}\right)=$$

$$=\lim_{n\to\infty}\frac{1-\frac{3}{n}}{\left(1+\frac{1}{n^2}\right)\left(3+\frac{1}{n}\right)}=\frac{1-0}{(1+0)(3+0)}=\frac{1}{3}.$$

3. Passage à la limite dans les inégalités.

Théorème 2.10. *Si à partir d'un certain rang ($n > N$) les termes d'une suite $\{x_n\}$ convergeant vers a sont tels que $x_n \geqslant b$* (resp. $x_n \leqslant b$), *alors $a \geqslant b$* (resp. $a \leqslant b$).

Démonstration. Supposons qu'à partir d'un certain rang ($n > N$) tous les termes x_n vérifient l'inégalité $x_n \geqslant b$. On demande de prouver que $a \geqslant b$. Supposons par absurde que $a < b$.

La suite $\{x_n\}$ convergeant vers a, pour $\varepsilon = b - a$ il existerait un entier N tel que pour $n > N$ on aurait $|x_n - a| < b - a$. Cette inégalité équivaut à la double inégalité : $-(b-a) < x_n - a < b - a$, d'où il résulte que $x_n < b$ pour $n > N$, ce qui est contraire à l'hypothèse. Donc $a \geqslant b$. Le cas $x_n \leqslant b$ se traite de façon analogue. ■

Corollaire 1. *Si à partir d'un certain rang ($n > N$) les termes de suites convergentes $\{x_n\}$ et $\{y_n\}$ sont tels que $x_n \leqslant y_n$, alors* $\lim\limits_{n\to\infty} x_n \leqslant \lim\limits_{n\to\infty} y_n$.

En effet, à partir d'un certain rang ($n > N$) les termes de la suite $\{y_n - x_n\}$ sont positifs, donc la limite de cette suite l'est aussi : $\lim\limits_{n\to\infty}(y_n - x_n) = \lim\limits_{n\to\infty} y_n - \lim\limits_{n\to\infty} x_n \geqslant 0$. D'où il s'ensuit que $\lim\limits_{n\to\infty} x_n \leqslant \lim\limits_{n\to\infty} y_n$.

Corollaire 2. *Si les termes d'une suite convergente $\{x_n\}$ appartiennent tous à un intervalle $[a, b]$, il en sera de même de la limite de cette suite.*

En effet, puisque $a \leqslant x_n \leqslant b$, on a $a \leqslant c \leqslant b$.

Le théorème suivant joue un rôle important dans les applications.

Théorème 2.11. *Soient données trois suites $\{x_n\}$, $\{y_n\}$ et $\{z_n\}$ telles que $x_n \leqslant y_n \leqslant z_n$ pour tout n et supposons que $\{x_n\}$ et $\{z_n\}$ admettent la même limite a. La suite $\{y_n\}$ admet aussi a pour limite.*

Démonstration. Considérons un $\varepsilon > 0$ arbitraire. Pour cet ε on peut exhiber un entier N_1 tel que $|x_n - a| < \varepsilon$

pour $n > N_1$, c'est-à-dire que

$$a - \varepsilon < x_n < a + \varepsilon. \tag{4}$$

Pour le même ε on peut trouver un entier N_2 tel que $|z_n - a| < \varepsilon$ pour $n > N_2$, c'est-à-dire que

$$a - \varepsilon < z_n < a + \varepsilon. \tag{5}$$

Soit $N = \max\{N_1, N_2\}$. Les inégalités (4) et (5) seront alors simultanément vérifiées pour $n > N$. En combinant ces inégalités à la double inégalité de l'hypothèse, on trouve

$$a - \varepsilon < x_n \leqslant y_n \leqslant z_n < a + \varepsilon \quad \text{pour} \quad n > N.$$

D'où

$$a - \varepsilon < y_n < a + \varepsilon \text{ ou } |y_n - a| < \varepsilon \quad \text{pour} \quad n > N.$$

Ce qui exprime que la suite $\{y_n\}$ admet a pour limite. ■

§ 3. Suites monotones

1. Définition et critère de convergence des suites monotones.

Définition. Une suite $\{x_n\}$ est dite *strictement croissante* si $x_n < x_{n+1}$; *croissante* si $x_n \leqslant x_{n+1}$; *strictement décroissante* si $x_n > x_{n+1}$; *décroissante* si $x_n \geqslant x_{n+1}$.

Toutes ces suites sont dites *monotones*. Les suites strictement croissantes et strictement décroissantes s'appellent *strictement monotones*.

Voyons des exemples de suites monotones.

1. La suite $1, 1/2, 1/3, \ldots, 1/n, \ldots$ est strictement décroissante et bornée.

2. La suite $1, 1, 1/2, 1/2, 1/3, \ldots, 1/n, 1/n, \ldots$ est décroissante et bornée.

3. La suite $1, 2, 3, \ldots, n, \ldots$, est strictement croissante et illimitée.

4. La suite $1, 1, 2, 2, 3, 3, \ldots, n, n, \ldots$ est croissante et illimitée.

5. La suite $1/2, 2/3, 3/4, \ldots, n/(n+1), \ldots$ est une suite strictement croissante et bornée.

Signalons que les suites monotones sont bornées au moins d'un côté: les suites croissantes, inférieurement ($x_n \geqslant x_1$), les suites décroissantes, supérieurement ($x_n \leqslant x_1$). Il s'avère que si une suite monotone est bornée (inférieurement et supérieurement), elle converge. Les suites non monotones ne jouissent pas de cette propriété. Par exemple, la suite non monotone $\{(-1)^n\}$ est bornée, mais non convergente (cf. remarque suivant le théorème 2.6).

On a le théorème fondamental suivant relativement aux suites monotones.

Théorème 2.12. *Une suite monotone bornée converge.*

Démonstration. Traitons le cas d'une suite croissante. Par hypothèse, $x_n \leqslant x_{n+1}$ et $x_n \leqslant M$ pour tout n. Considérons l'ensemble numérique X formé par les termes de cette suite. Cet ensemble est par hypothèse borné supérieurement et non vide. Donc il possède une borne supérieure d'après le théorème 1.1. Désignons cette borne par a et prouvons qu'elle est la limite de la suite $\{x_n\}$.

Le nombre a étant la borne supérieure de l'ensemble des termes de $\{x_n\}$, d'après la propriété de la borne supérieure pour tout $\varepsilon > 0$ on peut exhiber un entier N tel que $x_N > a - \varepsilon$. La suite $\{x_n\}$ étant croissante, on a $x_n > a - \varepsilon$ pour $n > N$. Par ailleurs, $x_n \leqslant a < a + \varepsilon$ par définition d'un majorant, c'est-à-dire que $|x_n - a| < \varepsilon$ pour $n > N$. Ce qui exprime que a est la limite de la suite $\{x_n\}$.

Le cas d'une suite décroissante se traite de façon analogue. ■

Remarque. Pour qu'une suite monotone soit bornée, il est nécessaire et suffisant qu'elle soit convergente.

En effet, si une suite monotone est bornée, elle converge en vertu du théorème 2.12; si une suite monotone converge, elle est bornée en vertu du théorème 2.6.

2. Le nombre *e*. Considérons la suite $\{x_n\}$ de terme général $x_n = \left(1 + \frac{1}{n}\right)^n$:

$$(1+1)^1,\ \left(1+\frac{1}{2}\right)^2,\ \ldots,\ \left(1+\frac{1}{n}\right)^n,\ \ldots$$

Montrons qu'elle converge. Pour cela il suffit de prouver que la suite $\{x_n\}$ est strictement croissante et majorée. En appliquant la formule du binôme de Newton *) (chap. VI, § 3, n° 4, formule (10)), on obtient

$$x_n = 1 + n\frac{1}{n} + \frac{n(n-1)}{2!}\frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!}\frac{1}{n^3} + \ldots$$

$$\ldots + \frac{n(n-1)(n-2)\ \ldots\ [n-(n-1)]}{n!}\frac{1}{n^n}.$$

Mettons cette expression sous la forme suivante :

$$x_n = 2 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \frac{1}{3!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \ldots$$

$$\ldots + \frac{1}{n!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ \ldots\ \left(1 - \frac{n-1}{n}\right). \qquad (1)$$

*) Newton Isaac (1642-1727), célèbre mathématicien, physicien, mécanicien et astronome anglais.

De façon analogue :

$$x_{n+1} = 2 + \frac{1}{2!}\left(1 - \frac{1}{n+1}\right) + \frac{1}{3!}\left(1 - \frac{1}{n+1}\right)\left(1 - \frac{2}{n+1}\right) + \ldots$$
$$\ldots + \frac{1}{(n+1)!}\left(1 - \frac{1}{n+1}\right)\left(1 - \frac{2}{n+1}\right) \ldots \left(1 - \frac{n}{n+1}\right).$$

On remarquera maintenant que $\left(1 - \frac{k}{n}\right) < \left(1 - \frac{k}{n+1}\right)$ pour $0 < k < n$. Donc chaque terme de l'expression de x_{n+1} est strictement supérieur au terme homologue de l'expression de x_n et en outre x_{n+1} contient un terme strictement positif de plus que x_n. Donc $x_n < x_{n+1}$, ce qui exprime que la suite x_n est strictement croissante.

Pour prouver qu'elle est majorée on remarquera que chaque parenthèse de l'expression (1) est strictement inférieure à 1. Comme $\frac{1}{n!} < \frac{1}{2^{n-1}}$ pour $n > 2$, on a

$$x_n < 2 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!} < 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}}.$$

La formule de la somme d'une progression géométrique nous conduit à l'inégalité

$$x_n < 1 + \frac{1 - 1/2^n}{1 - 1/2} = 3 - \frac{1}{2^{n-1}} < 3.$$

On a ainsi prouvé que la suite $\left\{\left(1 + \frac{1}{n}\right)^n\right\}$ est strictement croissante et majorée. D'après le théorème 2.12 elle admet une limite qui est désignée par la lettre e. Par définition donc

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n.$$

Signalons que le nombre e joue un grand rôle en mathématiques. Il est pris en particulier pour base des logarithmes népériens. Dans ce paragraphe on ne donne que la définition de e. Dans la suite on étudiera une méthode de calcul de e avec n'importe quelle précision.

Remarquons seulement qu'étant donné que $x_n < 3$ et que $2 < x_n$ (cf. formule (1)), on en déduit que $2 \leqslant e \leqslant 3$. On démontre que e est irrationnel.

Prouvons un théorème qui sera souvent utilisé pour la démonstration d'autres théorèmes.

§ 4. Théorème des segments emboîtés

Soit donnée une suite de segments $[a_1, b_1]$, $[a_2, b_2]$, ..., $[a_n, b_n]$, ... telle que $[a_1, b_1] \supset [a_2, b_2] \supset \ldots$ $\ldots \supset [a_n, b_n] \supset \ldots$, c'est-à-dire que

$$a_n \leqslant a_{n+1} < b_{n+1} \leqslant b_n \text{ pour tout } n. \qquad (1)$$

et supposons que $\lim_{n\to\infty}(b_n - a_n) = 0$. Cette suite s'appelle *suite de segments emboîtés.*

Théorème 2.13. *Il existe un seul point intérieur à tous les segments d'une suite de segments emboîtés.*

Démonstration. De (1) il s'ensuit que les extrémités gauches des segments forment une suite croissante

$$a_1 \leqslant a_2 \leqslant a_3 \leqslant \ldots \leqslant a_n \leqslant a_{n+1} \leqslant \ldots, \tag{2}$$

les extrémités droites, une suite décroissante

$$b_1 \geqslant b_2 \geqslant b_3 \geqslant \ldots \geqslant b_n \geqslant b_{n+1} \geqslant \ldots \tag{3}$$

De plus la suite (2) est majorée et la suite (3), minorée, puisque $a_n \leqslant b_1$ et $b_n \geqslant a_1$ pour tout n. Ces suites admettent donc des limites d'après le théorème 2.12. Soient $\lim_{n\to\infty} a_n = c'$ et $\lim_{n\to\infty} b_n = c''$. De la condition

$$\lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = c'' - c' = 0$$

il s'ensuit que $c' = c''$, c'est-à-dire que les suites $\{a_n\}$ et $\{b_n\}$ tendent vers la même limite. En désignant cette limite par c, on a la double inégalité $a_n \leqslant c \leqslant b_n$ pour tout n, autrement dit le point c est intérieur à tous les segments de la suite (1).

Prouvons maintenant que ce point est unique. Supposons qu'il existe un autre point $c_1 \neq c$ intérieur à tous les segments de la suite (1). Pour tout n on doit avoir $b_n - a_n \geqslant |c_1 - c|$ et par suite $\lim_{n\to\infty}(b_n - a_n) \geqslant |c_1 - c| \neq 0$, ce qui est contraire à l'hypothèse. ■

Remarque. Ce théorème est parfois mis en défaut si l'on remplace les segments par des intervalles ouverts. Ainsi les intervalles

$$]0, 1[\supset]0, 1/2[\supset]0, 1/4[\supset \ldots \supset]0, 1/2^n[\supset \ldots \tag{4}$$

ne possèdent pas de point qui leur soit commun à tous. En effet, quel que soit le point c pris sur l'intervalle $]0, 1[$, il existe toujours un entier N tel que pour $n > N$ l'on a $1/2^n < c$. Donc le point c ne sera pas contenu dans l'intervalle $]0, 1/2^{N+1}[$ et dans les suivants.

Pour la suite de l'exposé nous aurons besoin de certaines notions de géométrie analytique. Ces notions feront l'objet du chapitre III.

CHAPITRE III

GÉOMÉTRIE ANALYTIQUE DANS LE PLAN

La géométrie analytique est le domaine des mathématiques qui étudie les êtres géométriques par les méthodes algébriques. Le mathématicien français René Descartes a mis au point au XVII^{e} siècle la méthode des coordonnées qui est un important outil de géométrie analytique.

La méthode des coordonnées repose sur la notion de système de coordonnées. Nous étudierons les systèmes de coordonnées rectangulaires (ou cartésiennes) et polaires.

§ 1. Système de coordonnées rectangulaires

Deux axes Ox et Oy perpendiculaires se coupant en O et munis d'une même échelle (fig. 8) forment un *système de coordonnées rectangulaires dans le plan.*

Les axes Ox et Oy s'appellent *axes de coordonnées*; Ox est l'*axe des abscisses*; Oy, l'*axe des ordonnées*. Le plan contenant les axes Ox et Oy s'appelle *plan de coordonnées* et se note Oxy.

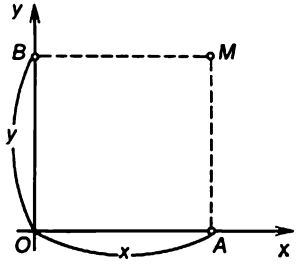


Fig. 8

Soit M un point quelconque du plan. Abaissons les perpendiculaires MA et MB sur les axes Ox et Oy.

On appellera *coordonnées rectangulaires* x et y du point M les mesures algébriques respectives $\overline{OA}$ et $\overline{OB}$: $x = \overline{OA}$, $y = \overline{OB}$.

Les coordonnées x et y du point sont respectivement son *abscisse* et son *ordonnée*. Le fait que M admet x et y pour coordonnées se note M $(x; y)$. A signaler que la première coordonnée est l'abscisse et la deuxième, l'ordonnée. L'origine des coordonnées a pour coordonnées (0; 0).

Donc dans le système de coordonnées choisi, à chaque point M du plan est associé un seul couple de nombres $(x; y)$ *) : ses coordonnées rectangulaires et, réciproquement, à chaque couple de nombres $(x; y)$ est associé un et un seul point M du plan Oxy d'abscisse x et d'ordonnée y.

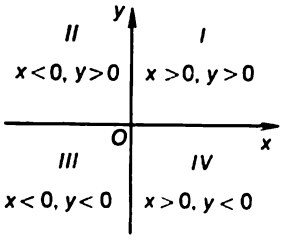


Fig. 9

Ainsi l'introduction d'un système de coordonnées rectangulaires dans le plan permet d'établir une correspondance bijective entre l'ensemble des points du plan et l'ensemble de tous les couples de nombres, d'où la possibilité d'appliquer les méthodes algébriques à la résolution des problèmes de géométrie.

Les axes de coordonnées divisent le plan en quatre parties appelées *quadrants* ou *angles de coordonnées*. Sur la figure 9 ces quadrants sont numérotés avec les chiffres romains I, II, III, IV.

§ 2. Problèmes élémentaires de géométrie analytique dans le plan

1. Distance de deux points.

Théorème 3.1. *La distance d de deux points quelconques $M_1(x_1; y_1)$ et $M_2(x_2; y_2)$ du plan est donnée par la formule*

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \quad (1)$$

Démonstration. Abaissons à partir des points M_1 et M_2 les perpendiculaires M_1B et M_2A respectivement sur les axes Oy et Ox et désignons par K le point d'intersection de M_1B et M_2A (fig. 10). Le point K a pour coordonnées $(x_2; y_1)$, donc (cf. chap. I, § 3)

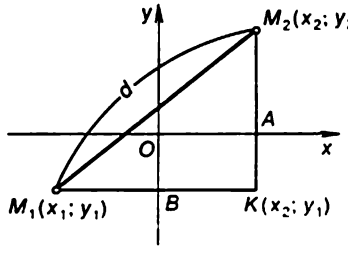


Fig. 10

$$M_1K = |x_2 - x_1|;$$
$$M_2K = |y_2 - y_1|.$$

Le triangle M_1M_2K étant rectangle, on a d'après le théorème de Pythagore

$$d = \sqrt{(M_1K)^2 + (M_2K)^2} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \quad \blacksquare$$

*) Nous avons affaire ici à un couple ordonné. Si $x \neq y$, les couples $(x; y)$ et $(y; x)$ sont distincts.

2. Aire d'un triangle.

Théorème 3.2. *L'aire s du triangle ABC formé par trois points* $A\ (x_1; y_1)$, $B\ (x_2; y_2)$ *et* $C\ (x_3; y_3)$ *non alignés est donnée par la formule*

$$s = \frac{1}{2}\,|[(x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)]|. \qquad (2)$$

Démonstration. L'aire du triangle ABC de la figure 11 est égale à

$$s_{ABC} = s_{ADEC} + s_{BCEF} - s_{ABFD}, \qquad (3)$$

où s_{ADEC}, s_{BCEF} et s_{ABFD} sont les aires des trapèzes $ADEC$, $BCEF$ et $ABFD$ respectivement. Comme

$$s_{ADEC} = DE\,\frac{AD+CE}{2} = \frac{(x_3 - x_1)(y_3 + y_1)}{2},$$

$$s_{BCEF} = EF\,\frac{EC+BF}{2} = \frac{(x_2 - x_3)(y_2 + y_3)}{2},$$

$$s_{ABFD} = DF\,\frac{AD+BF}{2} = \frac{(x_2 - x_1)(y_1 + y_2)}{2},$$

on obtient

$$s = \frac{1}{2}\,|[(x_1 - x_2)(y_1 + y_2) + (x_2 - x_3)(y_2 + y_3) + (x_3 - x_1)(y_3 + y_1)]|,$$

d'où l'on déduit la formule (2). La formule (2) se prouve de façon analogue pour toute autre disposition du triangle ABC. ■

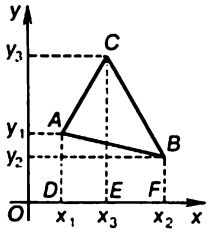


Fig. 11

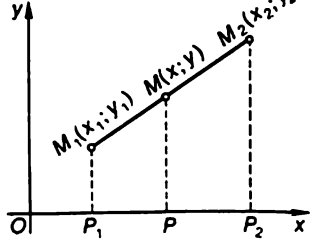


Fig. 12

Exemple. Soient donnés les points $A\ (1; 1)$, $B\ (6; 4)$, $C\ (8; 2)$. On demande de calculer l'aire du triangle ABC.

D'après la formule (2):

$$s = \frac{1}{2}\,|[(6-1)(2-1) - (8-1)(4-1)]| = \frac{1}{2}\,|[-16]| = 8.$$

3. Division d'un segment orienté dans un rapport donné. Soit donné dans le plan un segment orienté quelconque $\overrightarrow{M_1M_2}$ et soit M un point de $\overrightarrow{M_1M_2}$ distinct de M_2 (fig. 12).

Le nombre relatif λ défini par

$$\lambda = \frac{\overline{MM_1}}{\overline{MM_2}}, \tag{4}$$

s'appelle *rapport algébrique* dans lequel le point M divise le segment $\overrightarrow{M_1M_2}$.

Diviser un segment dans un rapport donné λ revient à déterminer les coordonnées du point M en fonction de λ et des coordonnées des points M_1 et M_2.

La solution de ce problème est donnée par le théorème suivant.

Théorème 3.3. *Si un point M $(x; y)$ divise un segment $\overrightarrow{M_1M_2}$ dans un rapport $\lambda \neq 1$, ses coordonnées sont définies par les formules*

$$x = \frac{\lambda x_2 - x_1}{\lambda - 1}; \quad y = \frac{\lambda y_2 - y_1}{\lambda - 1}, \tag{5}$$

où $(x_1; y_1)$ et $(x_2; y_2)$ sont les coordonnées respectives des points M_1 et M_2.

Démonstration. Supposons que la droite M_1M_2 n'est pas perpendiculaire à l'axe Ox. Soient P_1, P et P_2 les projetés respectifs des points M_1, M et M_2 sur l'axe Ox (fig. 12). D'après le théorème de Thalès

$$\frac{\overline{PP_1}}{\overline{PP_2}} = \frac{\overline{MM_1}}{\overline{MM_2}} = \lambda;$$

or $\overline{PP_1} = x_1 - x$, $\overline{PP_2} = x_2 - x$. Donc $\frac{x_1 - x}{x_2 - x} = \lambda$, d'où $x = \frac{\lambda x_2 - x_1}{\lambda - 1}$.

Si la droite M_1M_2 est perpendiculaire à Ox, on a $x_1 = x_2 = x$ et cette formule est également vraie. La deuxième formule (5) s'établit de façon analogue. ■

Corollaire. *Si M (x, y) est le milieu du segment $\overrightarrow{M_1M_2}$, alors $\lambda = -1$ et les formules (5) nous donnent*

$$x = \frac{x_1 + x_2}{2}, \quad y = \frac{y_1 + y_2}{2}.$$

Donc chaque coordonnée du milieu d'un segment est égale à la demi-somme des coordonnées respectives des extrémités de ce segment.

Exemple. Soient donnés les points M_1 (1; 1) et M_2 (7; 4). Trouver le point M $(x; y)$ qui divise le segment $\overrightarrow{M_1M_2}$ dans le rapport $\lambda = -1/2$.

Solution. En appliquant les formules (5) on trouve que les coordonnées de M sont: $x = 3$, $y = 2$.

§ 3. Coordonnées polaires

Le plus important après le système de coordonnées rectangulaires est le système de coordonnées polaires. Ce système est composé d'un point O, appelé *pôle*, et d'un axe OE issu de O, appelé *axe polaire*. On donne de plus une échelle pour mesurer la longueur des segments.

Soit donné un système de coordonnées polaires et soit M un point quelconque du plan. Soit ρ la distance de M à O; φ l'angle $(\overrightarrow{OE}, \overrightarrow{OM})$ (fig. 13).

On appelle *coordonnées polaires* du point M les nombres ρ et φ. Le nombre ρ qui est la première coordonnée s'appelle *rayon polaire*, le nombre φ, la seconde coordonnée, *angle polaire*.

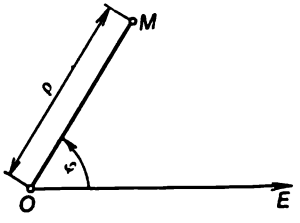


Fig. 13

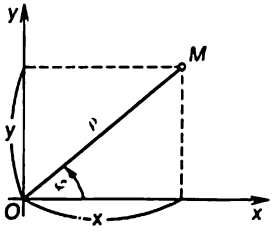


Fig. 14

Le point M de coordonnées polaires ρ et φ se note: $M(\rho; \varphi)$. Il est évident que le rayon polaire peut prendre toute valeur positive: $\rho \in [0, +\infty[$. On admet généralement que l'angle polaire varie dans l'intervalle $[0, 2\pi]$. Mais on est souvent amené à envisager des angles supérieurs à 2π ainsi que des angles négatifs.

Etablissons un lien entre les coordonnées polaires d'un point et ses coordonnées rectangulaires. On admettra que l'origine du système de coordonnées rectangulaires est confondue avec le pôle, et l'axe des abscisses, avec l'axe polaire. Supposons que le point M (fig. 14) admet x et y pour coordonnées rectangulaires et ρ et φ pour coordonnées polaires. Il est évident que

$$x = \rho \cos \varphi, \quad y = \rho \sin \varphi. \tag{1}$$

Les formules (1) expriment les coordonnées rectangulaires en fonction des coordonnées polaires. Les expressions des coordonnées polaires en fonction des coordonnées rectangulaires découlent des formules (1):

$$\rho = \sqrt{x^2 + y^2}, \quad \operatorname{tg} \varphi = y/x. \tag{2}$$

A noter que la formule $\operatorname{tg} \varphi = y/x$ nous donne deux valeurs de l'angle polaire φ, puisque φ varie entre 0 et 2π. De ces deux valeurs de φ on retient celle qui satisfait (1).

Exemple. Trouver les coordonnées polaires du point de coordonnées rectangulaires (2 ; 2) sachant que l'origine est confondue avec le pôle et l'axe des abscisses, avec l'axe polaire.

Solution. Les formules (2) nous donnent

$$\rho = 2\sqrt{2}, \quad \operatorname{tg}\varphi = 1.$$

De la deuxième égalité il résulte que $\varphi = \pi/4$ ou $\varphi = 5\pi/4$. Mais puisque $x = 2 > 0$ et $y = 2 > 0$, il faut retenir $\varphi = \pi/4$.

§ 4. Changement des coordonnées rectangulaires

Dans de nombreux problèmes de géométrie analytique on est amené à introduire deux systèmes de coordonnées rectangulaires. Il est évident que les coordonnées des points et les équations des courbes varient d'un système à l'autre. Le problème qui se pose est d'exprimer les coordonnées d'un point par rapport à un système en fonction des coordonnées de ce même point par rapport à l'autre système. Les formules de changement des coordonnées nous permettent de résoudre ce problème.

Considérons deux types de transformations des coordonnées rectangulaires :

1) *translation des axes de coordonnés*: l'origine des coordonnées varie mais la direction des axes reste fixe ;

2) *rotation des axes de coordonnées*: les deux axes tournent d'un même angle dans le même sens mais l'origine des coordonnées reste fixe.

1. Translation des axes. Supposons qu'un point M du plan a pour coordonnées $(x ; y)$ dans un système Oxy. Déplaçons l'origine des coordonnées en O' $(a ; b)$, où a et b sont les coordonnées de la nouvelle origine dans l'ancien système Oxy. Prenons les nouveaux axes de coordonnées $O'x'$ et $O'y'$ de même sens que les anciens. Désignons par $(x' ; y')$ les nouvelles coordonnées de M dans le système $O'x'y'$. Établissons une relation entre les nouvelles et les anciennes coordonnées de M. Projetons le point M sur les axes $O'x'$, $O'y'$ et Ox, Oy et le point O' sur les axes Ox et Oy. Soient $M_{x'}$, $M_{y'}$, M_x, M_y, O'_x et O'_y les projetés respectifs de ces points (fig. 15). La relation de Chasles (chap. I ; § 3) nous donne

$$x = \overline{OM_x} = \overline{OO'_x} + \overline{O'_xM_x} = \overline{OO'_x} + \overline{O'M_{x'}} = a + x',$$
$$y = \overline{OM_y} = \overline{OO'_y} + \overline{O'_yM_y} = \overline{OO'_y} + \overline{O'M_{y'}} = b + y'.$$

Ainsi

$$x = x' + a, \; y = y' + b \tag{1}$$

ou

$$x' = x - a, \; y' = y - b. \tag{2}$$

Ce que nous voulions.

2. **Rotation des axes de coordonnées.** Soit $Ox'y'$ le système de coordonnées obtenu par une rotation du système Oxy d'un angle α autour de O (fig. 16).

Supposons qu'un point M a $(x; y)$ pour coordonnées dans l'ancien système Oxy et $(x'; y')$ pour coordonnées dans le nouveau système $Ox'y'$. Etablissons une relation entre les anciennes et les

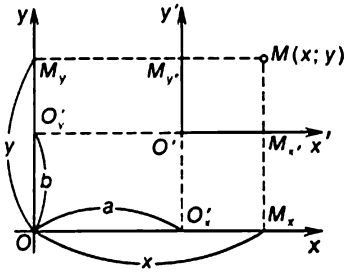


Fig. 15

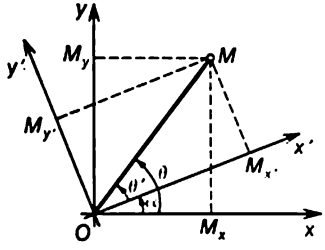


Fig. 16

nouvelles coordonnées du point M. A cet effet, appelons $(\rho; \theta)$ et $(\rho; \theta')$ les coordonnées polaires du point M dans les systèmes de pôle O et d'axes polaires Ox et Ox' respectivement. Il est évident que dans chaque cas $\rho = OM$ et $\theta = \theta' + \alpha$. Par ailleurs, d'après les formules (1) du § 3,

$$x = \rho \cos \theta, \quad y = \rho \sin \theta$$

et par analogie

$$x' = \rho \cos \theta', \quad y' = \rho \sin \theta'.$$

Donc

$$\begin{aligned} x = \rho \cos \theta &= \rho \cos (\theta' + \alpha) = \rho (\cos \theta' \cos \alpha - \\ &- \sin \theta' \sin \alpha) = \rho \cos \theta' \cos \alpha - \\ &- \rho \sin \theta' \sin \alpha = x' \cos \alpha - y' \sin \alpha; \\ y = \rho \sin \theta &= \rho \sin (\theta' + \alpha) = \rho (\cos \theta' \sin \alpha + \\ &+ \sin \theta' \cos \alpha) = \rho \cos \theta' \sin \alpha + \\ &+ \rho \sin \theta' \cos \alpha = x' \sin \alpha + y' \cos \alpha. \end{aligned}$$

Ainsi

$$\begin{cases} x = x' \cos \alpha - y' \sin \alpha, \\ y = x' \sin \alpha + y' \cos \alpha. \end{cases} \tag{3}$$

En exprimant x' et y' en fonction de x et y, on obtient

$$\begin{cases} x' = x \cos \alpha + y \sin \alpha, \\ y' = -x \sin \alpha + y \sin \alpha. \end{cases}$$

Exemple. Déterminer les coordonnées du point M (3 ; 5) dans un nouveau système de coordonnées $O'x'y'$ d'origine O' (— 2 ; 1) et d'axes parallèles à ceux de l'ancien système Oxy.

Solution. La formule (2) nous donne

$$x' = 3 + 2 = 5,\ y' = 5 - 1 = 4,$$

c'est-à-dire que les nouvelles coordonnées du point M sont (5 ; 4).

§ 5. Equation d'une courbe dans le plan

Soit une relation de la forme

$$F(x; y) = 0 \qquad (1)$$

entre les variables x et y. La relation (1) sera appelée *équation à deux variables x et y* si elle n'est vérifiée que pour certains couples $(x; y)$.

Exemples: $2x + 3y = 0$, $x^2 + y^2 - 25 = 0$, $\sin x + \sin y - 1 = 0$.

Si la relation (1) est valable pour tous les couples $(x; y)$ on l'appelle *identité*. Exemples:
$(x + y)^2 - x^2 - 2xy - y^2 = 0$, $(x + y)(x - y) - x^2 + y^2 = 0$.

La notion d'équation d'une courbe est très importante en géométrie analytique. Soit une courbe L dans un plan rapporté à un système de coordonnées rectangulaires (fig. 17).

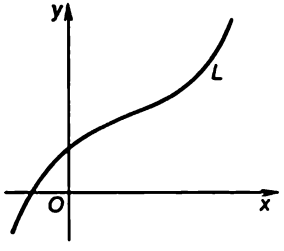


Fig. 17

Définition. L'équation (1) s'appelle *équation d'une courbe L* (dans un système de coordonnées donné) si elle est satisfaite par les coordonnées x et y de tous les points de L et de ces points seulement.

De cette définition il résulte que la courbe L est l'ensemble de tous les points du plan dont les coordonnées vérifient l'équation (1). On dira que l'équation (1) *définit* la courbe L.

La notion d'équation d'une courbe permet de résoudre les problèmes de géométrie par les méthodes algébriques. Par exemple, la recherche du point d'intersection des deux courbes d'équations $x + y = 0$ et $x^2 + y^2 = 1$ se ramène à la résolution d'un système de deux équations qui est un problème d'algèbre.

Une courbe L peut être définie par une équation de la forme

$$F(\rho; \varphi) = 0,$$

où $(\rho; \varphi)$ sont les coordonnées polaires du point courant.

Considérons des exemples d'équations de courbes.

1) $x - y = 0$. Cette équation s'écrit encore $y = x$. Elle est vérifiée par les coordonnées des points situés sur la bissectrice des quadrants I et III (fig. 18).

2) $x^2 - y^2 = 0$. Cette équation s'écrit encore $(x - y)(x + y) = 0$. L'ensemble des points dont les coordonnées vérifient cette équation est composé des bissectrices des quatre angles de coordonnées (fig. 19).

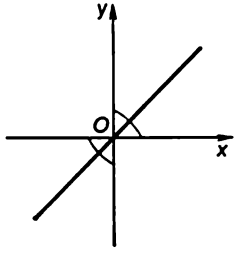


Fig. 18

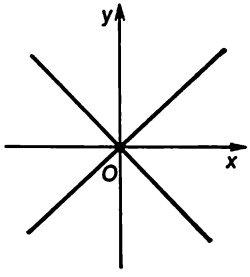


Fig. 19

3) $x^2 + y^2 = 0$. L'ensemble des points dont les coordonnées satisfont cette équation est composé du seul point (0 ; 0). On dit que cette équation définit une courbe dégénérée.

4) $x^2 + y^2 + 1 = 0$. Quels que soient x et y, on a toujours $x^2 + y^2 + 1 > 0$. Donc cette équation n'est vérifiée par les coordonnées d'aucun point, en d'autres termes elle ne définit aucune figure géométrique dans le plan.

5) $\rho = a \cos \varphi$, où $a > 0$, ρ et φ sont les coordonnées polaires du point courant M. Soit A le point de coordonnées polaires $(a ; 0)$ (fig. 20). Si $\rho = a \cos \varphi$, où $0 < \varphi < \pi/2$, l'angle $\widehat{OMA}$ est droit et réciproquement. Donc l'ensemble des points dont les coordonnées polaires vérifient cette équation est le cercle de diamètre OA.

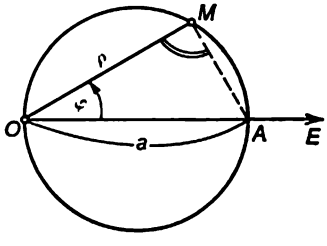


Fig. 20

6) $\rho = a\varphi$, où $a > 0$, ρ et φ les coordonnées polaires du point courant M. Si $\varphi = 0$, alors $\rho = 0$. Lorsque φ croît, le point M tourne autour du pôle O en s'éloignant simultanément de lui. L'ensemble des points dont les coordonnées polaires satisfont l'équation $\rho = a\varphi$ s'appelle *spirale d'Archimède* (fig. 21). On admet que φ peut prendre toute valeur positive.

Lorsque M accomplit un tour complet autour de O, l'angle φ croît de 2π et ρ, de $2a\pi$, c'est-à-dire que la spirale découpe sur

toute droite passant par le pôle des segments égaux (exception faite du segment contenant le pôle) de longueur $2a\pi$.

Dans les exemples ci-dessus on a étudié les propriétés d'une courbe d'après son équation et on a vu comment elle se présentait.

Considérons maintenant le problème inverse: trouver l'équation d'une courbe L dont les points jouissent de propriétés données.

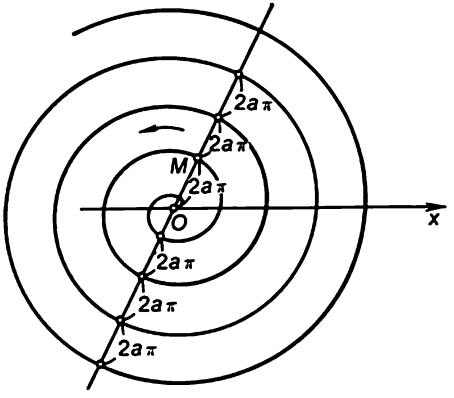


Fig. 21

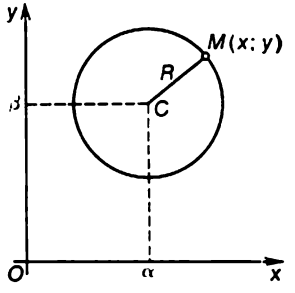


Fig. 22

Exemple. Etablir l'équation des points situés à une distance constante R d'un point fixe C $(\alpha; \beta)$. En d'autres termes, trouver l'équation du cercle de rayon R et de centre C $(\alpha; \beta)$ (fig. 22).

Solution. La distance entre un point quelconque M $(x; y)$ et C $(\alpha; \beta)$ est donnée par la formule

$$MC = \sqrt{(x - a)^2 + (y - \beta)^2}.$$

Si le point M est situé sur le cercle, c'est que $MC = R$ ou $MC^2 = R^2$, c'est-à-dire que les coordonnées de M vérifient l'équation

$$(x - \alpha)^2 + (y - \beta)^2 = R^2. \qquad (2)$$

Si le point M $(x; y)$ n'est pas sur le cercle, alors $MC^2 \neq R^2$, c'est-à-dire que les coordonnées de M ne satisfont pas l'équation (2). Donc l'équation du cercle cherchée est de la forme (2). Si l'on pose $\alpha = 0$ et $\beta = 0$ dans (2), on obtient l'équation du cercle de rayon R centré en O: $x^2 + y^2 = R^2$.

§ 6. Courbes du premier ordre

1. Equation d'une droite de coefficient directeur k. Soit donnée une droite L. Appelons *angle* de la droite L et de l'axe Ox l'angle α dont il faut faire pivoter l'axe Ox pour l'amener à coïncider avec L. L'angle α est défini à $\pm n\pi$, où n est un entier naturel. Le plus souvent pour angle α on prend le plus petit angle positif α dont il

faut faire pivoter l'axe Ox (dans le sens contraire aux aiguilles d'une montre) pour le faire coïncider avec la droite L (fig. 23). Dans ce cas $0 \leqslant \alpha < \pi$.

La tangente de l'angle de la droite L et de l'axe Ox s'appelle *coefficient directeur* ou *pente* de L et se désigne par la lettre k:

$$k = \operatorname{tg} \alpha. \tag{1}$$

De la formule (1) il résulte en particulier que si $\alpha = 0$, c'est-à-dire si la droite L est parallèle à Ox, alors $k = 0$. Si $\alpha = \pi/2$, c'est-à-dire si la droite L est perpendiculaire à Ox, $k = \operatorname{tg} \alpha$ n'a pas de sens. On dit alors que la droite L est de pente infinie.

Trouvons l'équation d'une droite dont on connaît la pente k et l'ordonnée à l'origine b *) (fig. 23).

Soit $M(x, y)$ un point arbitraire du plan. Si $k \neq 0$, le triangle BNM engendré par les droites BN et MN parallèles aux axes de coordonnées est rectangle. Pour que le point M soit situé sur la droite, il est nécessaire et suffisant que

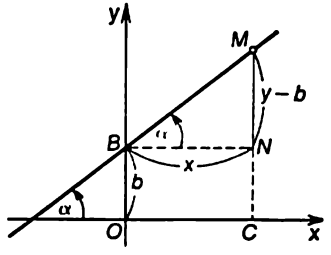


Fig. 23

$$\frac{\overline{NM}}{\overline{BN}} = \operatorname{tg} \alpha;$$

or $\overline{NM} = \overline{CM} - \overline{CN} = \overline{CM} - \overline{OB} = y - b$, $\overline{BN} = x$, donc une condition nécessaire et suffisante pour que le point $M(x; y)$ soit situé sur la droite est que

$$\frac{y-b}{x} = k. \tag{2}$$

Cette équation s'écrit encore

$$y = kx + b. \tag{3}$$

L'équation (3) s'appelle *équation d'une droite de coefficient directeur* k. Si $k = 0$, cette droite est parallèle à Ox et son équation devient $y = b$.

Ainsi toute droite non perpendiculaire à Ox admet une équation de la forme (3). La réciproque est visiblement vraie: toute équation de la forme (3) définit une droite de coefficient directeur k et d'ordonnée à l'origine b.

Exemple. Construire la droite d'équation $y = \frac{3}{4}x + 2$.

Solution. Repérons sur l'axe Oy le point B d'ordonnée 2 (fig. 24); traçons le segment BN de mesure algébrique $\overline{BN} = 4$

*) b est la mesure algébrique du segment orienté $\overrightarrow{OB}$.

parallèle à Ox puis le segment NM de mesure algébrique $\overline{NM} = 3$ parallèle à Oy. La droite BM est la droite cherchée. Son coefficient directeur est $k = \frac{3}{4}$ et son ordonnée à l'origine $b = 2$.

2. Equation d'une droite passant par un point donné et de coefficient directeur donné. On est souvent amené à former l'équation d'une droite dont on connaît un point $M_1\ (x_1\ ;\ y_1)$ et le coefficient directeur k. Ecrivons l'équation de cette droite sous la

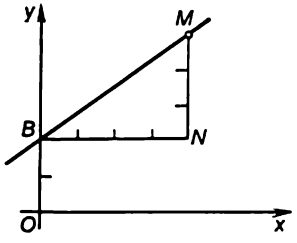


Fig. 24

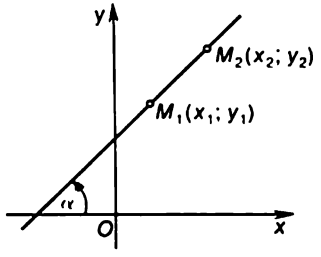


Fig. 25

forme (3), où b est un nombre provisoirement inconnu. Le point M_1 étant situé sur la droite, ses coordonnées x_1 et y_1 vérifient l'équation (3) : $y_1 = kx_1 + b$. En tirant b de cette relation et en le portant dans l'équation (3), on obtient l'équation cherchée :

$$y - y_1 = k\,(x - x_1). \tag{4}$$

Remarque. Si la droite passant par le point $M_1\ (x_1\ ;\ y_1)$ est perpendiculaire à l'axe Ox, c'est-à-dire si son coefficient directeur est infini, son équation est $x - x_1 = 0$. Formellement on peut déduire cette équation de (4) en divisant cette dernière par k et en faisant tendre k vers ∞.

3. Equation d'une droite passant par deux points donnés. Soient donnés deux points $M_1\ (x_1\ ;\ y_1)$ et $M_2\ (x_2\ ;\ y_2)$ (fig. 25). Mettons l'équation de la droite M_1M_2 sous la forme (4), où k, le coefficient directeur, est provisoirement inconnu. Le point M_2 étant sur la droite M_1M_2, ses coordonnées x_2 et y_2 satisfont l'équation (4) : $y_2 - y_1 = k\,(x_2 - x_1)$. En tirant k de cette équation (sous réserve que $x_1 \neq x_2$) et en le portant dans (4), on trouve l'équation cherchée :

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}\,(x - x_1).$$

Si $y_1 \neq y_2$, cette équation devient

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}. \tag{5}$$

Si $y_1 = y_2$, l'équation de la droite cherchée est de la forme $y = y_1$. Elle est alors parallèle à l'axe Ox. Si $x_1 = x_2$, la droite passant par les points M_1 et M_2 est parallèle à l'axe Oy et son équation est $x = x_1$.

Exemple. Former l'équation de la droite passant par les points $M_1(3; 1)$ et $M_2(5; 4)$.

Solution. En portant les coordonnées des points M_1 et M_2 dans la relation (5), on obtient l'équation cherchée :

$$\frac{x-3}{2} = \frac{y-1}{3} \text{ ou } 3x - 2y - 7 = 0.$$

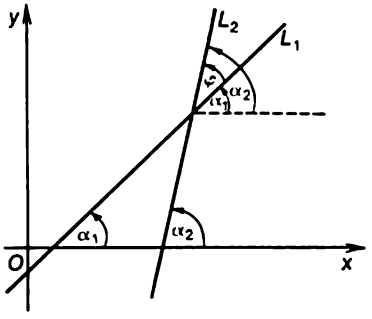


Fig. 26

4. Angles de deux droites. Soient deux droites L_1 et L_2 d'équations respectives $y = k_1x + b_1$, où $k_1 = \operatorname{tg} \alpha_1$, et $y = k_2x + b_2$, où $k_2 = \operatorname{tg} \alpha_2$. L'angle des droites L_1 et L_2 est l'ensemble $\{\varphi, \varphi + \pi\}$ (fig. 26).

Les angles α_1, α_2 et φ sont liés par la relation : $\alpha_2 = \alpha_1 + \varphi$ ou $\varphi = \alpha_2 - \alpha_1$. On a

$$\operatorname{tg} \varphi = \operatorname{tg} (\alpha_2 - \alpha_1) = \frac{\operatorname{tg} \alpha_2 - \operatorname{tg} \alpha_1}{1 + \operatorname{tg} \alpha_1 \operatorname{tg} \alpha_2},$$

ou

$$\operatorname{tg} \varphi = \frac{k_2 - k_1}{1 + k_2 k_1}. \tag{6}$$

Exemple. Trouver l'angle des droites d'équations $y = 2x + 3$ et $y = -3x + 2$.

Solution. Il est évident que $k_1 = 2$ et $k_2 = -3$. Donc d'après la formule (6),

$$\operatorname{tg} \varphi = (-3 - 2)/(1 + (-3) \cdot 2) = -5/-5 = 1.$$

L'un des angles est donc égal à $\pi/4$ et l'autre à $\pi - \pi/4 = 3\pi/4$.

5. Conditions de parallélisme et de perpendicularité de deux droites. Si les droites L_1 et L_2 sont parallèles, $\varphi = 0$ et $\operatorname{tg} \varphi = 0$. Le numérateur de la formule (6) est donc égal à 0 : $k_2 - k_1 = 0$ d'où

$$k_2 = k_1.$$

Donc, *pour que deux droites soient parallèles, il faut et il suffit que leurs coefficients directeurs soient égaux.*

Si les droites L_1 et L_2 sont perpendiculaires, c'est-à-dire si $\varphi = \pi/2$, alors $\alpha_2 = \pi/2 + \alpha_1$, $\operatorname{tg} \alpha_2 = \operatorname{tg} (\pi/2 + \alpha_1) =$

$= -\operatorname{cotg}\alpha_1 = -1/\operatorname{tg}\alpha_1$, en d'autres termes

$$k_2 = -\frac{1}{k_1}$$

ou encore

$$k_1 k_2 = -1.$$

Donc, *pour que deux droites soient perpendiculaires, il faut et il suffit que le produit de leurs coefficients directeurs soit égal à* -1.

On peut obtenir la condition nécessaire à partir de la formule (6). En effet, si $k_1 k_2 = -1$, le dénominateur de (6) est nul et $\operatorname{tg}\varphi = \infty$, donc $\varphi = \pi/2$.

6. Equation générale de la droite.

Théorème 3.4. *Dans le plan rapporté à un système de coordonnées rectangulaires Oxy toute droite est définie par une équation du premier degré*

$$Ax + By + C = 0, \quad (7)$$

et réciproquement, toute équation (7) avec des coefficients A, B et C arbitraires (A et B sont non tous deux nuls) définit une droite.

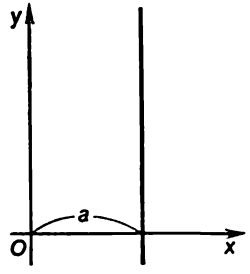


Fig. 27

Démonstration. Prouvons la première proposition du théorème. Au n° 1 on a vu que si une droite n'est pas perpendiculaire à l'axe Ox, son équation est $y = kx + b$. On reconnaît ici une équation (7) dans laquelle $A = k$, $B = -1$, et $C = b$. Si la droite est perpendiculaire à l'axe Ox, tous ses points ont la même abscisse a (fig. 27). L'équation de cette droite est $x = a$, c'est-à-dire est aussi une équation de la forme (7) avec $A = 1$, $B = 0$ et $C = -a$. Ce que nous voulions. Prouvons maintenant la réciproque. Soit donnée une équation (7) avec des coefficients A et B non tous deux nuls.

Si $B \neq 0$, on peut mettre (7) sous la forme

$$y = -\frac{A}{B}x - \frac{C}{B}.$$

En posant $k = -A/B$ et $b = -C/B$, on obtient l'équation $y = kx + b$, c'est-à-dire une équation de la forme (3) qui définit une droite.

Si $B = 0$, alors $A \neq 0$ et (7) devient $x = -C/A$. En désignant $-C/A$ par a, on obtient $x = a$, c'est-à-dire l'équation d'une droite perpendiculaire à l'axe Ox. ■

Les courbes définies dans un système de coordonnées rectangulaires par une équation du premier degré s'appellent *courbes du*

premier ordre. Donc toute droite est une courbe du premier ordre et inversement, toute courbe du premier ordre est une droite.

L'équation $Ax + By + C = 0$ s'appelle *équation générale de la droite*. Elle contient tous les cas particuliers de l'équation d'une droite.

7. Equation incomplète du premier degré. Equation de la droite en fonction de ses coordonnées à l'origine. Considérons trois cas particuliers où l'équation $Ax + By + C = 0$ est *incomplète*, c'est-à-dire que l'un des coefficients est nul.

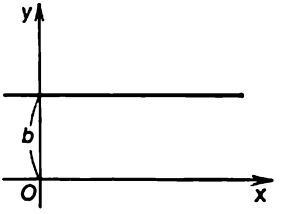


Fig. 28

1) $C = 0$; l'équation $Ax + By = 0$ définit une droite passant par l'origine des coordonnées.

2) $B = 0$ $(A \neq 0)$: l'équation $Ax + C = 0$ définit une droite parallèle à l'axe Oy. On a montré dans le théorème 3.4 que cette équation se ramène à la forme $x = a$, où $a = -C/A$ est l'abscisse à l'origine (cf. fig. 27). En particulier, si $a = 0$, la droite est confondue avec l'axe Oy. Donc l'équation $x = 0$ définit l'axe des ordonnées.

3) $A = 0$ $(B \neq 0)$: l'équation $By + C = 0$ définit une droite parallèle à l'axe Ox. Ce fait s'établit comme dans le cas précédent. Si l'on pose $-C/B = b$, l'équation devient $y = b$, où b est l'ordonnée à l'origine (fig. 28). En particulier, si $b = 0$, la droite est confondue avec l'axe Ox. Donc l'équation $y = 0$ définit l'axe des abscisses.

Soit donnée maintenant l'équation $Ax + By + C = 0$ avec des coefficients A, B et C non nuls. Mettons-la sous la forme

$$\frac{x}{-C/A} + \frac{y}{-C/B} = 1.$$

En posant $a = -C/A$ et $b = -C/B$, on obtient

$$\frac{x}{a} + \frac{y}{b} = 1. \tag{8}$$

L'équation (8) s'appelle *équation de la droite en fonction des coordonnées à l'origine*. Les nombres a et b sont respectivement l'abscisse et l'ordonnée à l'origine. Cette forme d'écriture est commode pour la construction de la droite.

Exemple. Soit la droite d'équation $3x - 5y + 15 = 0$. Représenter cette équation en fonction des coordonnées à l'origine et construire cette droite.

S o l u t i o n. L'équation de cette droite en fonction des coordonnées à l'origine s'écrit

$$\frac{x}{-5}+\frac{y}{3}=1.$$

Repérons sur l'axe Ox le point M_1 d'abscisse -5 et sur l'axe Oy le point M_2 d'ordonnée 3 et traçons la droite M_1M_2 (fig. 29).

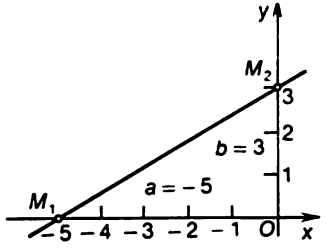


Fig. 29

8. Position relative de deux droites.

Soient L_1 et L_2 des droites d'équations

$$A_1x + B_1y + C_1 = 0,$$
$$A_2x + B_2y + C_2 = 0.$$

Traitons ces équations comme un système d'équations du premier degré à deux inconnues x et y. La solution de ce système est

$$x=\frac{B_1C_2-B_2C_1}{A_1B_2-A_2B_1}\ ; \qquad y=\frac{A_2C_1-A_1C_2}{A_1B_2-A_2B_1}\ .$$

Supposons que $A_1B_2-A_2B_1 \neq 0$. Les formules ci-dessus nous donnent alors la solution du système (s'en assurer !). Ceci exprime que les droites L_1 et L_2 se coupent au point de coordonnées x et y.

Supposons maintenant que $A_1B_2 - A_2B_1 = 0$. Distinguons deux cas:

1) $A_2C_1 - A_1C_2 = 0$ et $B_1C_2 - B_2C_1 = 0$.

2) $A_2C_1 - A_1C_2 \neq 0$ $(B_1C_2 - B_2C_1 \neq 0)$.

Dans le cas 1) on a $A_2 = \mu A_1$, $B_2 = \mu B_1$, $C_2 = \mu C_1$ ou

$$\frac{A_2}{A_1}=\frac{B_2}{B_1}=\frac{C_2}{C_1}=\mu,$$

où $\mu \neq 0$ est une constante. Ceci exprime que les coefficients des équations sont proportionnels, donc que la deuxième équation est le produit de la première par μ. Dans ce cas les droites L_1 et L_2 sont confondues, c'est-à-dire que ces équations représentent la même droite. Il est évident que le système admet une infinité de solutions.

Dans le cas 2), si par exemple $A_2C_1 - A_1C_2 \neq 0$, alors en supposant que le système admet une solution $(x_0; y_0)$, on obtiendrait une contradiction. En effet, en substituant x_0 et y_0 à x et y, en multipliant la première équation par A_2 et la deuxième par A_1, on obtient $A_2C_1 - A_1C_2 = 0$, ce qui contredit l'hypothèse. Donc le

système ne possède pas de solutions. Dans ce cas les droites L_1 et L_2 ne se rencontrent pas, c'est-à-dire sont parallèles.

Ainsi deux droites du plan sont soit concourantes, soit confondues, soit parallèles.

9. Equation normale de la droite. Distance d'un point à une droite. Soit donnée une droite L. Menons par l'origine des coordonnées une droite n perpendiculaire à L et appelons-la normale à

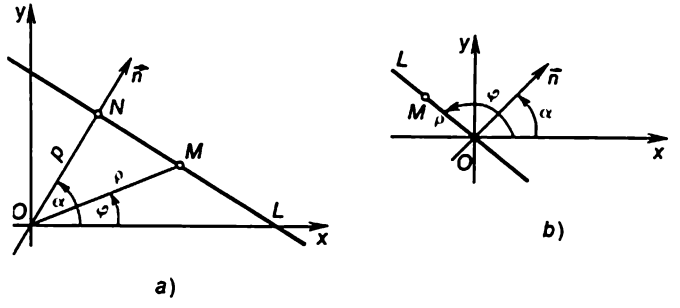


Fig. 30

L. Soit N le pied de n (fig. 30, *a*). Orientons n dans le sens de O vers N. Si N et O sont confondus, on prendra l'un quelconque des deux sens.

Désignons par α l'angle $(\overrightarrow{Ox}, \overrightarrow{On})$ et par p la longueur de ON. Donc $0 \leqslant \alpha < 2\pi$, $p \geqslant 0$. Etablissons l'équation de la droite L en admettant que α et p sont connus. A cet effet, prenons sur la droite L un point quelconque M de coordonnées polaires $(\rho; \varphi)$, où O est le pôle et Ox, l'axe polaire. Si O et N ne sont pas confondus, on déduit du triangle rectangle ONM que

$$p = \rho \cos(\alpha - \varphi) = \rho(\cos\alpha \cos\varphi + \sin\alpha \sin\varphi).$$

Cette égalité s'écrit encore

$$\rho \cos\varphi \cos\alpha + \rho \sin\varphi \sin\alpha - p = 0. \qquad (9)$$

Vu que les points extérieurs à la droite L ne satisfont pas l'équation (9), cette dernière est l'équation de la droite L en coordonnées polaires. D'après les formules de passage des coordonnées rectangulaires aux coordonnées polaires, on a : $\rho \cos\varphi = x$, $\rho \sin\varphi = y$. Donc dans un système de coordonnées rectangulaires l'équation (9) s'écrit

$$x \cos\alpha + y \sin\alpha - p = 0. \qquad (10)$$

Si les points O et N sont confondus, la droite L passe par l'origine des coordonnées (fig. 30, *b*) et $p = 0$. Dans ce cas il est évident que pour tout point M de L on a $\cos(\varphi - \alpha) = 0$. En multipliant

par ρ on obtient $\rho \cos(\varphi - \alpha) = 0$, d'où $\rho \cos\varphi \cos\alpha + \rho \sin\varphi \sin\alpha = 0$ ou

$$x \cos\alpha + y \sin\alpha = 0.$$

Donc dans ce cas aussi l'équation de la droite est de la forme (10).

L'équation (10) s'appelle *équation normale* de la droite L.

L'équation normale de la droite permet de déterminer la distance d'un point donné du plan à une droite.

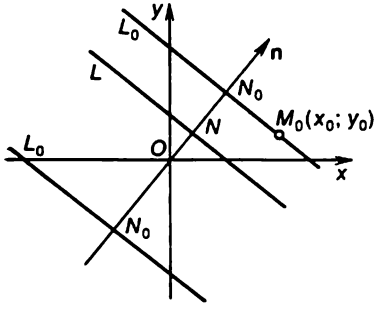


Fig. 31

Soit L une droite d'équation $x \cos\alpha + y \sin\alpha - p = 0$ et soit $M_0(x_0; y_0)$ un point extérieur à L. On demande de calculer la distance d du point M_0 à la droite L.

Menons par M_0 une droite L_0 parallèle à L. Soient N_0 le point d'intersection de L_0 et de la normale, p_0 la longueur de ON_0 (fig. 31).

Si les points N et N_0 sont situés d'un même côté de O, l'équation normale de la droite L_0 est $x \cos\alpha + y \sin\alpha - p_0 = 0$. Puisque $M_0(x_0; y_0) \in L_0$, on a $x_0 \cos\alpha + y_0 \sin\alpha - p_0 = 0$, d'où $p_0 = x_0 \cos\alpha + y_0 \sin\alpha$. Dans ce cas

$$d = |p_0 - p| = |x_0 \cos\alpha + y_0 \sin\alpha - p|.$$

Si les points N et N_0 sont situés de part et d'autre de O, l'équation normale de L_0 est $x \cos\alpha_1 + y \sin\alpha_1 - p_0 = 0$, où α_1 est supplémentaire de α. Donc $p_0 = x_0 \cos\alpha_1 + y_0 \sin\alpha_1 = -x_0 \cos\alpha - y_0 \sin\alpha$. Dans ce cas

$$d = |p_0 + p| = |-x_0 \cos\alpha - y_0 \sin\alpha + p| = |x_0 \cos\alpha + y_0 \sin\alpha - p|.$$

Dans chacun des deux cas on obtient la formule

$$d = |x_0 \cos\alpha + y_0 \sin\alpha - p|. \qquad (11)$$

Signalons que la formule (11) est valable dans le cas aussi où le point $M_0(x_0; y_0)$ est situé sur la droite L, c'est-à-dire que ses coordonnées vérifient l'équation de la droite L: $x_0 \cos\alpha + y_0 \sin\alpha - p = 0$. La formule (11) nous donne alors $d = 0$. A l'examen de la formule (11), on voit que pour calculer la distance d d'un point M_0 à la droite L, il faut porter les coordonnées x_0 et y_0 de M_0 dans le premier membre de l'équation normale de L et prendre le module de l'expression obtenue.

Voyons maintenant comment on passe de l'équation générale à l'équation normale d'une droite. Soit

$$Ax + By + C = 0 \tag{12}$$

l'équation générale d'une droite,

$$x \cos \alpha + y \sin \alpha - p = 0, \tag{13}$$

l'équation normale.

Les équations (12) et (13) définissant la même droite, leurs coefficients sont proportionnels. En multipliant les deux membres de l'équation (12) par un facteur $\mu \neq 0$, on obtient l'équation

$$\mu Ax + \mu By + \mu C = 0.$$

Si le facteur μ est convenablement choisi, cette équation se transforme en l'équation (13). Pour cela il faut que

$$\mu A = \cos \alpha, \quad \mu B = \sin \alpha \text{ et } \mu C = -p. \tag{14}$$

Elevons les deux premières égalités au carré et ajoutons-les. On obtient

$$\mu^2 (A^2 + B^2) = \cos^2 \alpha + \sin^2 \alpha = 1,$$

d'où

$$\mu = \pm \frac{1}{\sqrt{A^2 + B^2}}. \tag{15}$$

Le nombre μ s'appelle *facteur de normalisation*. Son signe se détermine à partir de la troisième égalité (14). D'après cette égalité μC est strictement négatif si $C \neq 0$. Donc dans la formule (15) on prend le signe contraire à celui de C. Si $C = 0$, le signe de μ peut être arbitrairement choisi.

Ainsi pour ramener l'équation générale à l'équation normale il faut trouver le facteur de normalisation μ et ensuite multiplier tous les termes de l'équation générale par μ.

Exemple. Trouver la distance d du point M (4 ; 3) à la droite $3x - 4y + 10 = 0$.

Solution. Ramenons cette équation à la forme normale. Le facteur de normalisation vaut :

$$\mu = -1/\sqrt{3^2 + 4^2} = -1/5.$$

En multipliant l'équation donnée par μ, on obtient l'équation normale

$$-\frac{3}{5} x + \frac{4}{5} y - 2 = 0.$$

La formule (11) nous donne la distance d cherchée :

$$d = |(-3/5) \cdot 4 + (4/5) \cdot 3 - 2| = |-2| = 2.$$

§ 7. Courbes du second ordre

Considérons trois types de courbes : l'ellipse, l'hyperbole et la parabole qui sont définies par des équations du second degré dans un système de coordonnées rectangulaires. De telles courbes s'appellent *courbes du second ordre* ou encore *coniques*.

1. Ellipse.

Définition. On appelle *ellipse* l'ensemble des points du plan dont la somme des distances à deux points fixes, appelés *foyers*, est une constante strictement supérieure à la distance des foyers.

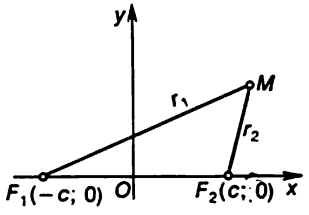


Fig. 32

Désignons les foyers par F_1 et F_2, la distance F_1F_2 par $2c$ et la somme des distances d'un point quelconque de l'ellipse aux foyers, par $2a$. Par définition, $2a > 2c$ ou $a > c$.

Considérons un système de coordonnées rectangulaires dont l'axe des abscisses passe par les foyers F_1 et F_2 et dont l'origine est située au milieu de F_1F_2. Les coordonnées des foyers sont alors : $F_1\ (-c;\ 0)$, $F_2\ (c;\ 0)$ (fig. 32). Trouvons l'équation de l'ellipse dans ce système de coordonnées.

Soit $M\ (x;\ y)$ un point arbitraire du plan. Désignons par r_1 et r_2 les distances respectives du point M aux foyers ($r_1 = F_1M$, $r_2 = F_2M$). Les nombres r_1 et r_2 sont les *rayons vecteurs* du point M. Par définition, pour que le point M appartienne à l'ellipse considérée, il est nécessaire et suffisant que

$$r_1 + r_2 = 2a. \tag{1}$$

La formule (1) du § 2 nous donne

$$r_1 = \sqrt{(x+c)^2+y^2},\quad r_2 = \sqrt{(x-c)^2+y^2}. \tag{2}$$

En portant ces expressions dans l'égalité (1), on trouve

$$\sqrt{(x+c)^2+y^2} + \sqrt{(x-c)^2+y^2} = 2a. \tag{3}$$

L'équation (3) est l'équation cherchée. Mais cette équation est peu commode à l'usage, c'est pourquoi on la ramène à une forme plus simple. Faisons passer le deuxième radical à droite et élevons les deux membres de l'équation au carré :

$$(x+c)^2+y^2 = 4a^2 - 4a\sqrt{(x-c)^2+y^2} + (x-c)^2+y^2,$$

ou

$$a\sqrt{(x-c)^2+y^2} = a^2 - cx. \tag{4}$$

Elevons encore les deux membres au carré:

$$a^2x^2 - 2a^2cx + a^2c^2 + a^2y^2 = a^4 - 2a^2cx + c^2x^2,$$

d'où

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2). \tag{5}$$

Introduisons la nouvelle quantité

$$b = \sqrt{a^2 - c^2}, \tag{6}$$

dont la signification géométrique apparaîtra plus loin. Vu que $a > c$ par hypothèse, $a^2 - c^2 > 0$ et par suite $b > 0$. De (6) on déduit que

$$b^2 = a^2 - c^2.$$

Donc l'équation (5) peut être mise sous la forme

$$b^2x^2 + a^2y^2 = a^2b^2.$$

En divisant les deux membres par a^2b^2, on obtient en définitive

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{7}$$

Puisque l'équation (7) résulte de l'équation (3), les coordonnées de tout point de l'ellipse vérifiant l'équation (3) vérifieront l'équation (7). Mais en ramenant l'équation (3) à une forme plus simple on l'a élevée deux fois au carré, ce qui a pu introduire des racines étrangères et violer l'équivalence des équations (7) et (3). Assurons-nous que si les coordonnées d'un point satisfont l'équation (7), elles satisferont aussi l'équation (3), autrement dit montrons que ces équations sont équivalentes. Pour cela il suffit de toute évidence de montrer que les rayons vecteurs r_1 et r_2 de tout point dont les coordonnées vérifient l'équation (7) vérifient la relation (1). En effet, supposons que les coordonnées x et y d'un point satisfont l'équation (7). En portant dans l'expression (2) la quantité $y^2 = b^2\left(1 - \frac{x^2}{a^2}\right)$ déduite de (7), on trouve après quelques transformations simples que $r_1 = \sqrt{(a + cx/a)^2}$. Puisque $|x| \leqslant a$ (ceci découle de (7)) et $c/a < 1$, on a $a + cx/a > 0$ et donc $r_1 = a + cx/a$.

On trouve de façon analogue que $r_2 = a - cx/a$. En ajoutant ces égalités terme à terme, on obtient la relation (1). Ce que nous voulions. Donc tout point dont les coordonnées vérifient l'équation (7) est situé sur l'ellipse et réciproquement, ce qui exprime que l'équation (7) est l'équation de l'ellipse. L'équation (7) s'appelle *équation canonique de l'ellipse.* Donc l'ellipse est une conique.

Etudions maintenant la forme de l'ellipse d'après son équation canonique (7). Remarquons que l'équation (7) ne contient que des termes en x et y de puissances paires, donc l'ellipse est symétri-

que par rapport aux axes Ox et Oy et par rapport à l'origine des coordonnées. Pour connaître la forme complète de l'ellipse, il nous suffit de la connaître dans le premier quadrant seulement.

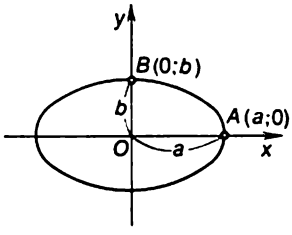


Fig. 33

Dans cette région $y \geqslant 0$, donc en résolvant l'équation (7) par rapport à y, on trouve

$$y = \frac{b}{a}\sqrt{a^2 - x^2}. \qquad (8)$$

La relation (8) entraîne les propositions suivantes.

1) Si $x = 0$, $y = b$. Donc le point $(0; b)$ que l'on désignera par B est situé sur l'ellipse.

2) y décroît lorsque x croît de 0 à a.

3) Si $x = a$, $y = 0$. Donc le point $(a; 0)$ que l'on désignera par A se trouve sur l'ellipse.

4) Si $x > a$, les valeurs de y sont imaginaires. Donc les points tels que $x > a$ n'appartiennent pas à l'ellipse.

Ainsi la partie de l'ellipse située dans le premier quadrant est l'arc BA *) (fig. 33).

On obtient l'ellipse tout entière par une symétrie par rapport aux axes de coordonnées.

Remarque. Si $a = b$, l'équation (7) devient $x^2 + y^2 = a^2$. C'est l'équation d'un cercle de rayon a. Donc le cercle est un cas particulier de l'ellipse. Remarquons que l'ellipse est la transformée soit de son cercle principal (le cercle de centre O et de rayon a) dans une affinité orthogonale d'axe Ox et de rapport $\frac{b}{a}$, soit de son cercle secondaire (le cercle de centre O et de rayon b) dans une affinité orthogonale d'axe Oy et de rapport $\frac{a}{b}$. Dans la première affinité tout point $(x; y)$ se transforme en un point $(x; y_1)$, où $y_1 = \frac{b}{a}y$. En portant $y = \frac{a}{b}y_1$ dans l'équation du cercle on trouve l'équation de l'ellipse

$$\frac{x^2}{a^2} + \frac{(y_1)^2}{b^2} = 1.$$

Les axes de symétrie de l'ellipse sont appelés *axes de l'ellipse*, et le centre de symétrie (le point d'intersection des axes), *centre de l'ellipse*. Les points d'intersection de l'ellipse et des axes sont les *sommets de l'ellipse*. L'ellipse découpe sur les axes des segments égaux à $2a$ et $2b$. De l'égalité (6) il s'ensuit que $a \geqslant b$. Les quantités a et b s'appellent respectivement *demi-grand axe* et *demi-petit axe*. Les axes respectifs s'appellent *grand* et *petit axe*.

*) Au chapitre VI on introduira la notion de *convexité* d'une courbe $y = f(x)$ et on montrera que l'arc BA est convexe vers le haut.

Introduisons encore une quantité qui caractérise la forme de l'ellipse.

Définition. On appelle *excentricité* de l'ellipse le rapport $\frac{c}{a}$, où c est la moitié de la distance focale, a le demi-grand axe de l'ellipse.

L'excentricité se désigne par $\varepsilon = \frac{c}{a}$. Puisque $c < a$, on a $0 \leqslant \varepsilon < 1$, c'est-à-dire que l'excentricité de l'ellipse est strictement inférieure à 1. Puisque $c^2 = a^2 - b^2$, on trouve

$$\varepsilon^2 = \frac{c^2}{a^2} = \frac{a^2 - b^2}{a^2} = 1 - \left(\frac{b}{a}\right)^2,$$

d'où

$$\frac{b}{a} = \sqrt{1 - \varepsilon^2}.$$

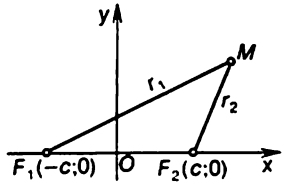


Fig. 34

Cette égalité nous permet de donner une interprétation géométrique de l'excentricité de l'ellipse. Lorsque ε est très petit, les nombres a et b sont presque égaux et l'ellipse est voisine du cercle. Si ε est voisin de 1, le nombre b est très petit en regard de a et l'ellipse est fortement allongée le long du grand axe. Donc l'excentricité de l'ellipse caractérise son allongement.

On sait que les planètes et certaines comètes se déplacent suivant des trajectoires elliptiques. Mais les excentricités des orbites des planètes sont très petites, tandis que celles des orbites des comètes sont voisines de l'unité. Donc les planètes se déplacent sur des trajectoires presque circulaires, alors que les comètes tantôt se rapprochent du Soleil (qui occupe un foyer), tantôt s'en éloignent considérablement.

2. Hyperbole.

Définition. On appelle *hyperbole* l'ensemble des points du plan dont le module de la différence des distances à deux points fixes, appelés foyers, est une quantité constante strictement inférieure à la distance focale.

Désignons les foyers par F_1 et F_2, la distance focale F_1F_2, par $2c$ et le module de la différence des distances aux foyers, par $2a$. Par définition, $2a < 2c$ ou $a < c$.

Pour former l'équation de l'hyperbole ramenons le plan à un système de coordonnées rectangulaires dont l'axe des abscisses passe par les foyers F_1 et F_2 et dont l'origine est le milieu de de F_1F_2. Les foyers de l'hyperbole ont alors pour coordonnées: $F_1\ (-c;\ 0)$, $F_2\ (c;\ 0)$ (fig. 34). Soit $M\ (x;\ y)$ un point quelconque du plan. MF_1 et MF_2 sont les rayons vecteurs du point M. Désignons-les par r_1 et r_2. Par définition, le point $M\ (x;\ y)$ sera situé

sur l'hyperbole considérée si et seulement si $| r_1 - r_2 | = 2a$. D'où

$$r_1 - r_2 = \pm 2a. \tag{9}$$

La formule (1) du § 2 nous donne

$$r_1 = \sqrt{(x+c)^2 + y^2}, \quad r_2 = \sqrt{(x-c)^2 + y^2}. \tag{10}$$

En portant ces expressions dans (9), on obtient

$$\sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = \pm 2a. \tag{11}$$

L'équation (11) est l'équation cherchée de l'hyperbole. Réduisons cette équation à une forme plus simple comme nous l'avons fait pour l'équation (3) de l'ellipse. Portons le deuxième radical dans le second membre, puis élevons au carré. On obtient

$$(x+c)^2 + y^2 = 4a^2 \pm 4a\sqrt{(x-c)^2 + y^2} + (x-c)^2 + y^2$$

ou

$$cx - a^2 = \pm a\sqrt{(x-c)^2 + y^2}. \tag{12}$$

En élevant encore au carré, on obtient

$$c^2x^2 - 2a^2cx + a^4 = a^2x^2 - 2a^2cx + a^2c^2 + a^2y^2.$$

D'où

$$(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2). \tag{13}$$

Considérons la quantité

$$b = \sqrt{c^2 - a^2} \tag{14}$$

dont la signification géométrique sera précisée plus loin. Puisque $c > a$, on a $c^2 - a^2 > 0$ et $b > 0$. L'égalité (14) entraîne

$$b^2 = c^2 - a^2.$$

L'équation (13) devient

$$b^2x^2 - a^2y^2 = a^2b^2$$

ou

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{15}$$

Comme pour l'ellipse on démontre l'équivalence des équations (15) et (11). L'équation (15) s'appelle *équation canonique de l'hyperbole.*

Etudions la forme de l'hyperbole à l'aide de son équation canonique. L'équation (15) ne contenant que des puissances paires en x et y, l'hyperbole est symétrique par rapport aux axes Ox et Oy et par rapport à l'origine des coordonnées. On peut donc limiter notre étude à la portion d'hyperbole comprise dans le premier quadrant. Dans cette région du plan $y \geqslant 0$, donc la résolu-

tion de l'équation (15) par rapport à y nous donne

$$y = \frac{b}{a}\sqrt{x^2 - a^2}. \tag{16}$$

L'égalité (16) entraîne les propositions suivantes.

1) Si $0 \leqslant x < a$, y prend des valeurs imaginaires. Donc les points d'abscisse $0 \leqslant x < a$ n'appartiennent pas à l'hyperbole.

2) Si $x = a$, $y = 0$, donc le point $A\ (a; 0)$ est sur l'hyperbole.

3) Si $x > a$, $y > 0$ et de plus $y \to +\infty$ lorsque $x \to +\infty$. Le point courant $M\ (x; y)$ se déplace sur l'hyperbole à partir du

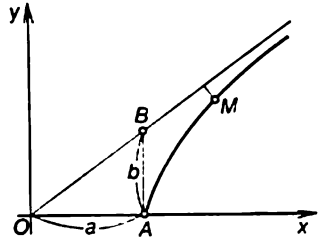


Fig. 35

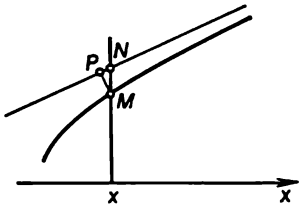


Fig. 36

point $A\ (a; 0)$ vers la « droite » et en « haut » lorsque x croît (fig. 35). Précisons comment le point M « s'éloigne à l'infini ».

A cet effet considérons l'équation

$$y = \frac{b}{a}x \tag{17}$$

qui définit une droite de coefficient directeur $k = \frac{b}{a}$, passant par l'origine des coordonnées. La portion de cette droite située dans le premier quadrant est représentée sur la figure 35. Cette droite est confondue avec l'hypoténuse OB du triangle rectangle OAB de côtés $OA = a$ et $AB = b$.

Montrons que tout en s'éloignant à l'infini, le point M se rapproche indéfiniment de la droite (17) qui s'appelle *asymptote* de l'hyperbole *).

Soient x une valeur arbitraire $(x \geqslant a)$ et $M\ (x; y)$ et $N\ (x; Y)$ deux points tels que

$$y = \frac{b}{a}\sqrt{x^2 - a^2} \text{ et } Y = \frac{b}{a}x.$$

Le point M est situé sur l'hyperbole, le point N, sur la droite (17). Ces deux points ayant la même abscisse x, la droite MN est perpendiculaire à l'axe Ox (fig. 36). Trouvons la longueur du segment MN.

*) Au chap. VI on donnera la définition de l'asymptote du graphique d'une fonction $y = f(x)$ et on montrera que la droite $y = \frac{b}{a}x$ est une asymptote de l'hyperbole. On étudiera *ibidem* le sens de la convexité de l'hyperbole.

Remarquons tout d'abord que pour $x \geqslant a$

$$Y = \frac{b}{a} x = \frac{b}{a} \sqrt{x^2} > \frac{b}{a} \sqrt{x^2 - a^2} = y.$$

Ce qui exprime que le point de l'hyperbole d'abscisse x se trouve en dessous du point d'abscisse x de l'asymptote. Donc

$$MN = Y - y = \frac{b}{a} x - \frac{b}{a} \sqrt{x^2 - a^2} = \frac{b}{a} (x - \sqrt{x^2 - a^2}) =$$

$$= \frac{b}{a} \frac{(x - \sqrt{x^2 - a^2})(x + \sqrt{x^2 - a^2})}{x + \sqrt{x^2 - a^2}} = \frac{ab}{x + \sqrt{x^2 - a^2}}.$$

De cette expression il s'ensuit que MN tend vers 0 lorsque $x \to +\infty$, puisque le numérateur ab est constant et le dénominateur tend vers $+\infty$.

Soit P le pied de la perpendiculaire abaissée du point M sur la droite (17). MP est alors la distance de M à cette droite. Il

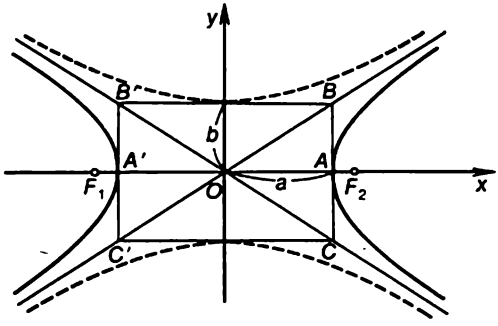


Fig. 37

est évident que $MP < MN$ et puisque $MN \to 0$, à fortiori $MP \to 0$ lorsque $x \to +\infty$, c'est-à-dire que le point M se rapproche indéfiniment de la droite (17). C.Q.F.D.

On peut construire maintenant l'hyperbole tout entière par symétrie par rapport aux axes de coordonnées (fig. 37). L'hyperbole est composée de deux branches (droite et gauche) et admet deux asymptotes : $y = \frac{b}{a} x$ et $y = -\frac{b}{a} x$. La première asymptote a déjà été étudiée, la seconde est symétrique de la première par rapport à l'axe Ox (ou Oy).

Les axes de symétrie s'appellent *axes* de l'hyperbole, le centre de symétrie (le point d'intersection des axes), *centre* de l'hyperbole. L'un des axes coupe l'hyperbole en deux points appelés *sommets* (les points A' et A sur la figure 37). Cet axe est l'*axe transverse* de l'hyperbole. L'autre axe n'a pas de points communs avec l'hyperbole. On l'appelle *axe non transverse*. Le rectangle

$BB'C'C$ de côtés $2a$ et $2b$ (fig. 37) s'appelle *rectangle principal* de l'hyperbole. Les quantités a et b s'appellent respectivement *demi-axe transverse* et *demi-axe non transverse* de l'hyperbole.

L'équation

$$\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1$$

définit aussi une hyperbole, qui est représentée en pointillé sur la figure 37. Ses sommets sont situés sur l'axe Oy. Cette hyperbole est dite *conjuguée* de l'hyperbole (15). Ces deux hyperboles possèdent les mêmes asymptotes.

Une hyperbole dont les demi-axes sont égaux ($a = b$) s'appelle *équilatère*. Son équation canonique est

$$x^2 - y^2 = a^2.$$

Les asymptotes d'une hyperbole équilatère sont perpendiculaires, puisque son rectangle principal est un carré.

Définition. On appelle *excentricité* de l'hyperbole le rapport $\frac{c}{a}$, où c est la moitié de la distance focale et a, le demi-axe transverse de l'hyperbole.

L'excentricité de l'hyperbole se désigne aussi par ε. L'excentricité de l'hyperbole est > 1, puisque $c > a$. En remarquant que $c^2 = a^2 + b^2$, on trouve

$$\varepsilon^2 = \frac{c^2}{a^2} = \frac{a^2+b^2}{a^2} = 1 + \left(\frac{b}{a}\right)^2,$$

d'où

$$\frac{b}{a} = \sqrt{\varepsilon^2 - 1}.$$

La dernière égalité nous permet de donner une interprétation géométrique de l'excentricité de l'hyperbole. Plus l'excentricité est petite, c'est-à-dire plus elle est proche de 1, plus le rapport $\frac{b}{a}$ est petit. Ce qui exprime que le rectangle principal est plus allongé le long de l'axe transverse. Donc l'excentricité de l'hyperbole caractérise la forme du rectangle principal et par conséquent la forme même de l'hyperbole.

Dans le cas d'une hyperbole équilatère, $\varepsilon = \sqrt{2}$.

3. Directrices de l'ellipse et de l'hyperbole.

Définition 1. On appelle *directrices de l'ellipse* les deux droites perpendiculaires au grand axe aux points d'abscisses $\pm \frac{a}{\varepsilon}$.

Les équations des directrices d'une ellipse définie par (7) sont :

$$x = -\frac{a}{\varepsilon} \text{ et } x = \frac{a}{\varepsilon}.$$

Puisque pour une ellipse $\varepsilon < 1$, on a $\frac{a}{\varepsilon} > a$. Donc les directrices sont situées à l'extérieur de l'ellipse (fig. 38).

Définition 2. On appelle *directrices de l'hyperbole* les deux droites perpendiculaires à l'axe transverse aux points d'abscisses $\pm \frac{a}{\varepsilon}$.

Les équations des directrices d'une hyperbole définie par l'équation canonique (15) sont

$$x = -\frac{a}{\varepsilon} \text{ et } x = \frac{a}{\varepsilon}.$$

Puisque pour une hyperbole $\varepsilon > 1$, on a $\frac{a}{\varepsilon} < a$. Donc une directrice est située entre le centre et le sommet droit de l'hyperbole, l'autre directrice, entre le centre et le sommet gauche (fig. 39).

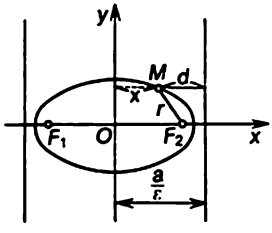


Fig. 38

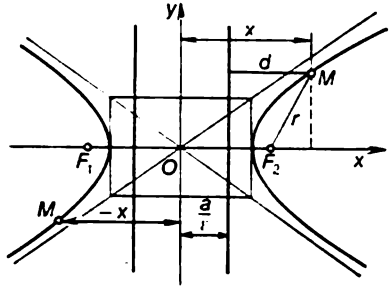


Fig. 39

Les notions de directrice et d'excentricité permettent de formuler une propriété générale inhérente à l'ellipse et à l'hyperbole. On a les deux théorèmes suivants.

Théorème 3.5. *Le rapport $\frac{r}{d}$ du rayon vecteur r d'un point M de l'ellipse à la distace d de ce point à la directrice associée est une constante égale à l'excentricité de l'ellipse.*

Démonstration. Supposons pour fixer les idées qu'il est question du rayon vecteur MF_2 et de la directrice associée au foyer F_2. La distance d du point M $(x; y)$ à cette directrice est

$$d = \frac{a}{\varepsilon} - x \tag{18}$$

(fig. 38). Les égalités (2) et (4) entraînent

$$r = r_2 = \sqrt{(x-c)^2 + y^2} = a - \frac{c}{a}x.$$

En posant $\frac{c}{a} = \varepsilon$, on obtient la formule de la distance de M au foyer F_2:

$$r = a - \varepsilon x. \tag{19}$$

Des relations (18) et (19) on déduit que

$$\frac{r}{d} = \frac{a - \varepsilon x}{\frac{a}{\varepsilon} - x} = \frac{(a - \varepsilon x)\varepsilon}{a - \varepsilon x} = \varepsilon. \blacksquare$$

Théorème 3.6. *Le rapport $\frac{r}{d}$ du rayon vecteur d'un point M de l'hyperbole à la distance d de ce point à la directrice associée est une constante égale à l'excentricité de l'hyperbole.*

Démonstration. Supposons pour fixer les idées qu'il s'agit du rayon vecteur MF_2 et de la directrice associée au foyer F_2. Soit $M(x; y)$ un point arbitraire de l'hyperbole (fig. 39). Considérons les deux cas suivants.

1) Le point M est situé sur la branche droite de l'hyperbole. La distance d du point M à la directrice droite est

$$d = x - \frac{a}{\varepsilon}. \tag{20}$$

Les égalités (10) et (12) nous donnent

$$r = r_2 = \sqrt{(x-c)^2 + y^2} = \frac{c}{a}x - a.$$

En posant $\frac{c}{a} = \varepsilon$, on obtient la formule du rayon vecteur MF_2

$$r = \varepsilon x - a. \tag{21}$$

Les relations (20) et (21) entraînent

$$\frac{r}{d} = \frac{\varepsilon x - a}{x - \frac{a}{\varepsilon}} = \frac{(\varepsilon x - a)\varepsilon}{\varepsilon x - a} = \varepsilon.$$

2) Le point M est situé sur la branche gauche de l'hyperbole. La distance d du point M à la directrice droite vaut (fig. 39)

$$d = -x + \frac{a}{\varepsilon}. \tag{22}$$

Par analogie à (21), le rayon vecteur MF_2 est égal à

$$r = -\left(\frac{c}{a}x - a\right) = -(\varepsilon x - a). \tag{23}$$

Les relations (22) et (23) nous donnent

$$\frac{r}{d} = \frac{-(\varepsilon x - a)}{-x + \frac{a}{\varepsilon}} = \frac{(-\varepsilon x + a)\varepsilon}{(-\varepsilon x + a)} = \varepsilon. \blacksquare$$

Cette propriété de l'ellipse et de l'hyperbole peut servir de définition de ces courbes : *l'ensemble des points dont le rapport des distances au foyer et à la directrice associée est une constante* ε *est une ellipse si* $\varepsilon < 1$ *et une hyperbole si* $\varepsilon > 1$.

Une question se pose immédiatement : quelle est la nature de cet ensemble lorsque $\varepsilon = 1$. Il s'avère que c'est une conique appelée parabole.

4. Parabole.

Définition. On appelle *parabole* l'ensemble des points du plan équidistants d'un point fixe appelé *foyer* et d'une droite fixe ne passant pas par le foyer, appelée *directrice*.

Pour établir l'équation de la parabole, rapportons le plan à un système de coordonnées rectangulaires dont l'axe des abscisses passe par le foyer perpendiculairement à la directrice, prenons pour sens positif le sens qui va de la directrice au foyer et plaçons l'origine à égale distance du foyer et de la directrice.

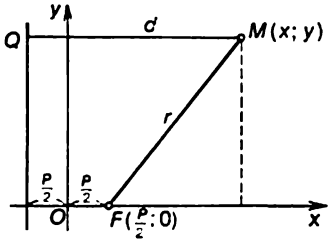


Fig. 40

Soit $M(x\,;\,y)$ un point du plan. Désignons par r la distance du point M au foyer F, par d, la distance de M à la directrice et par p, la distance du foyer à la directrice (fig. 40). La quantité p s'appelle *paramètre* de la parabole ; sa signification géométrique apparaîtra plus bas. Pour que le point M appartienne à la parabole il est nécessaire et suffisant que

$$r = d. \tag{24}$$

Le foyer F a pour coordonnées $(p/2\,;\,0)$; donc la formule (1) du § 2 nous donne

$$r = FM = \sqrt{(x - p/2)^2 + y^2}. \tag{25}$$

La distance d s'exprime visiblement par l'égalité (fig. 40)

$$d = MQ = x + \frac{p}{2}. \tag{26}$$

Signalons que cette formule n'est valable que pour $x \geqslant 0$. Si $x < 0$, il est évident que $r > d$ et par conséquent le point $M\,(x\,;\,y)$ n'est pas situé sur la parabole. En portant dans (24) les expressions (25) et (26) de r et de d, on trouve

$$\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = x + \frac{p}{2}. \tag{27}$$

Ceci est l'équation cherchée de la parabole. Ramenons-la à une forme plus commode. En élevant au carré, on obtient

$$x^2 - px + p^2/4 + y^2 = x^2 + px + p^2/4$$

ou

$$y^2 = 2px. \tag{28}$$

Voyons si l'élévation au carré n'a pas introduit de racines étrangères. Pour cela il suffit de montrer que la relation (28) est réalisée pour tout point $M(x; y)$ dont les coordonnées satisfont l'équation (28). En effet, de l'équation (28) il résulte que $x \geqslant 0$, donc $d = p/2 + x$ pour les points $M(x; y)$ d'abscisses positives. En portant l'expression (28) de y^2 dans (25) et puisque $x \geqslant 0$, on obtient $r = p/2 + x$, c'est-à-dire que $r = d$. Ce que nous voulions. Donc l'équation (28) est satisfaite par les coordonnées de tous les points de la parabole et de ces points seulement, ce qui exprime que l'équation (28) est l'équation de la parabole considérée.

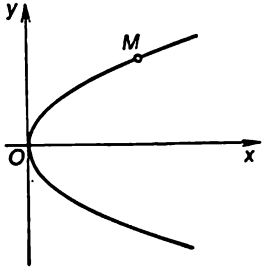


Fig. 41

L'équation (28) s'appelle *équation canonique de la parabole*. Cette équation étant du second degré, la parabole est une conique.

Etudions maintenant la forme de la parabole à l'aide de son équation (28). La parabole est symétrique par rapport à l'axe Ox puisque y figure avec un exposant pair dans l'équation (28). Il suffit donc de l'étudier dans le demi-plan supérieur. Dans cette région $y \geqslant 0$, donc la résolution de l'équation (28) par rapport à y nous donne

$$y = \sqrt{2px}. \tag{29}$$

Cette égalité entraîne les assertions suivantes :

1) Si $x < 0$, y prend des valeurs imaginaires. Donc aucun point de la parabole n'est situé à gauche de l'axe Oy, ce que nous avons déjà signalé.

2) Si $x = 0$, $y = 0$. Donc l'origine des coordonnées appartient à la parabole.

3) y croît avec x et de plus $y \to +\infty$ lorsque $x \to +\infty$.

Donc le point courant $M(x; y)$ se déplace à partir de l'origine « à droite » et « vers le haut » et de plus il s'éloigne indéfiniment de l'axe Oy et de l'axe Ox lorsque $x \to +\infty$.

Une symétrie par rapport à l'axe Ox nous donne la parabole tout entière (fig. 41).

Le point O s'appelle *sommet* de la parabole, l'axe de symétrie (l'axe Ox), *axe* de la parabole. Le paramètre p représente la distan-

ce du foyer à la directrice. Voyons comment le paramètre influe sur la forme de la parabole. A cet effet, prenons une valeur bien définie de l'abscisse, par exemple $x = 1$, et trouvons les valeurs correspondantes de l'ordonnée à l'aide de l'équation (28) : $y = \pm \sqrt{2p}$. On obtient deux points $M_1(1; +\sqrt{2p})$ et $M_2(1; -\sqrt{2p})$ symétriques par rapport à l'axe et distants de $2\sqrt{2p}$.

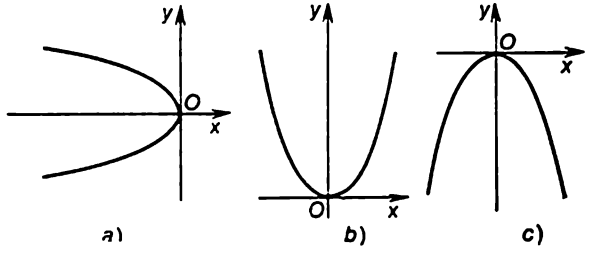


Fig. 42

On remarque que cette distance est directement proportionnelle à p. Donc le paramètre p caractérise la « largeur » du domaine borné par la parabole.

La parabole d'équation $y^2 = -2px$, $p > 0$, est située à gauche de l'axe Oy (fig. 42, *a*). Elle a pour sommet l'origine des coordonnées et pour axe de symétrie, l'axe Ox.

L'équation $x^2 = 2py$, $p > 0$, est l'équation d'une parabole dont le sommet est l'origine des coordonnées et l'axe de symétrie, l'axe Oy (fig. 42, *b*). Cette parabole est située au-dessus de l'axe des abscisses. L'équation $x^2 = -2py$, $p > 0$, définit une parabole située au-dessous de l'axe Ox et ayant l'origine des coordonnées pour sommet (fig. 42, *c*).

§ 8. Equation générale des coniques

Un important problème de géométrie analytique est l'étude de l'équation générale des coniques et leur réduction aux formes canoniques.

L'équation générale d'une conique est

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0, \qquad (1)$$

où les coefficients A, $2B$, C, $2D$, $2E$ et F *) sont des nombres quelconques, et en outre A, B et C sont non tous nuls, c'est-à-dire que $A^2 + B^2 + C^2 \neq 0$.

*) Pour faciliter les transformations de l'équation (1) on a désigné les coefficients de xy, x et y respectivement par $2B$, $2D$ et $2E$.

1. Réduction de l'équation générale d'une conique à la forme canonique.

Lemme 3.1. *Soit donnée l'équation* (1) *dans un système de coordonnées rectangulaires Oxy et supposons que* $AC - B^2 \neq 0$. *Par le produit d'une translation et d'une rotation des axes de coordonnées l'équation* (1) *se réduit à la forme*

$$A'x''^2 + C'y''^2 + F' = 0, \tag{2}$$

où A', C' *et* F' *sont des nombres*; $(x''; y'')$ *les nouvelles coordonnées.*

Démonstration. Supposons que le système Oxy se transforme par une translation en un système $O'x'y'$ dont l'origine O' a pour coordonnées $(x_0; y_0)$ dans le système Oxy. Les anciennes coordonnées $(x; y)$ sont liées aux nouvelles $(x'; y')$ par les formules

$$x = x' + x_0, \quad y = y' + y_0$$

(cf. formules (1), § 4). L'équation (1) devient dans les nouvelles coordonnées

$$Ax'^2 + 2Bx'y' + Cy'^2 + 2D'x' + 2E'y' + F' = 0, \tag{3}$$

où

$$D' = Ax_0 + By_0 + D; \quad E' = Bx_0 + Cy_0 + E;$$
$$F' = Ax_0^2 + 2Bx_0y_0 + Cy_0^2 + 2Dx_0 + 2Ey_0 + F.$$

Les coefficients D' et E' s'annulent si l'on choisit les coordonnées du point $(x_0; y_0)$ telles que

$$\begin{cases} Ax_0 + By_0 + D = 0, \\ Bx_0 + Cy_0 + E = 0. \end{cases} \tag{4}$$

Le système (4) admet une seule solution, puisque $AC - B^2 \neq 0$.

Si un couple de nombres $(x_0; y_0)$ est solution du système (4), l'équation (3) devient

$$Ax'^2 + 2Bx'y' + Cy'^2 + F' = 0. \tag{5}$$

Supposons maintenant que le système $O'x'y'$ se transforme par une rotation d'angle α en le système $O'x''y''$. Les coordonnées x', y' sont liées aux coordonnées x'', y'' par les formules

$$x' = x'' \cos\alpha - y'' \sin\alpha, \quad y' = x'' \sin\alpha + y'' \cos\alpha$$

(cf. formules (3), § 4). L'équation (5) s'écrit dans le système $O'x''y''$

$$A'x''^2 + 2B'x''y'' + C'y''^2 + F' = 0, \tag{6}$$

où

$$A' = A\cos^2\alpha + 2B\cos\alpha\sin\alpha + C\sin^2\alpha;$$
$$B' = -A\sin\alpha\cos\alpha + B(\cos^2\alpha - \sin^2\alpha) + C\sin\alpha\cos\alpha;$$
$$C' = A\sin^2\alpha - 2B\cos\alpha\sin\alpha + C\cos^2\alpha.$$

Choisissons l'angle α de manière à annuler le coefficient B' de l'équation (6). Ceci nous conduit à l'équation $2B \cos 2\alpha = (A - C) \sin 2\alpha$ par rapport à α. Si $A = C$, $\cos 2\alpha = 0$ et l'on peut poser $\alpha = \pi/4$. Si $A \neq C$, on prend $\alpha = \frac{1}{2} \operatorname{Arc\,tg} \frac{2B}{A-C}$ et l'équation (6) devient

$$A'x''^2 + C'y''^2 + F' = 0,$$

c'est-à-dire qu'on a obtenu l'équation (2). ■

Remarque. Les équations (4) s'appellent *équations du centre de la conique*, le point $(x_0; y_0)$, où x_0, y_0 représentent la solution du système (4), est le *centre* de cette conique. A noter qu'une condition nécessaire et suffisante pour que le système (4) admette une seule solution est que le nombre $AC - B^2$, appelé déterminant du système (4) (chap. 10, § 2), soit non nul.

2. Invariance de l'expression $AC - B^2$. Classification des coniques. Les coefficients A, B et C de l'équation (1) ne changent pas par une translation des axes de coordonnées ainsi qu'il résulte de la démonstration du lemme 3.1. par contre, ils sont modifiés par une rotation. Cependant l'expression $AC - B^2$ est invariante aussi bien par une translation que par une rotation du système de coordonnées, en d'autres termes elle ne dépend pas de la transformation du système de coordonnées. En effet, ceci est évident pour la translation [cf. formules (1) et (5)] ; vérifions-le pour la rotation. Utilisons pour cela les expressions des coefficients A', B' et C' de l'équation (6). On a

$$\begin{aligned} A'C' - B'^2 = {} & (A \cos^2 \alpha + 2B \sin \alpha \cos \alpha + C \sin^2 \alpha) \times \\ & \times (A \sin^2 \alpha - 2B \sin \alpha \cos \alpha + C \cos^2 \alpha) - \\ & - [(C - A) \sin \alpha \cos \alpha + B (\cos^2 \alpha - \sin^2 \alpha)]^2. \end{aligned}$$

En chassant les parenthèses et en réduisant les termes semblables on obtient

$$\begin{aligned} A'C' - B'^2 = {} & AC (\cos^2 \alpha + \sin^2 \alpha)^2 - \\ & - B^2 (\cos^2 \alpha + \sin^2 \alpha)^2 = AC - B^2, \end{aligned}$$

C.Q.F.D.

La quantité $AC - B^2$ s'appelle *invariant* de l'équation générale des coniques. Cet invariant est capital dans l'étude des coniques.

Une conique est du genre

1) ellipse si $AC - B^2 > 0$;
2) hyperbole si $AC - B^2 < 0$;
3) parabole si $AC - B^2 = 0$.

Considérons les courbes des trois genres.

1) *Genre ellipse.* Puisque $AC - B^2 > 0$, l'équation générale peut être en vertu du lemme 3.1 réduite à la forme (pour la commodité de l'écriture on omettra les accents des coefficients et des coordonnées):

$$Ax^2 + Cy^2 + F = 0.$$

Distinguons les cas suivants:

a) $A > 0$, $C > 0$ (le cas $A < 0$, $C < 0$ se ramène au cas $A > 0$, $C > 0$ par une multiplication de l'équation par -1) et $F < 0$. Faisons passer F dans le second membre et divisons par F. L'équation devient

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

où $a^2 = -F/A$, $b^2 = -F/C$. On reconnaît ici l'équation canonique de l'ellipse.

b) $A > 0$, $C > 0$, $F > 0$. En procédant comme dans a), on obtient l'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1$$

qui n'est satisfaite par aucun point du plan. On l'appelle *équation de l'ellipse imaginaire.*

c) $A > 0$, $C > 0$, $F = 0$. L'équation devient ($a^2 = A$, $c^2 = C$)

$$a^2x^2 + c^2y^2 = 0.$$

Elle n'est vérifiée que par les coordonnées du point (0, 0). Cette équation définit un *couple de droites imaginaires concourantes.*

2) *Genre hyperbole.* Puisque $AC - B^2 < 0$, l'équation générale devient en vertu du lemme 3.1

$$Ax^2 + Cy^2 + F = 0.$$

Distinguons les cas suivants:

a) $A > 0$, $C < 0$ (le cas $A < 0$, $C > 0$ se ramène au cas $A > 0$, $C < 0$ par une multiplication de l'équation par -1) et $F \neq 0$. Supposons par exemple que $F < 0$. En faisant passer F à droite et en divisant par lui, on obtient l'équation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

où $a^2 = -F/A$, $b^2 = F/C$, qui est l'équation canonique de l'hyperbole.

b) $A > 0$, $C < 0$, $F = 0$. L'équation devient ($a^2 = A$, $c^2 = -C$):

$$a^2x^2 - c^2y^2 = 0 \text{ ou } (ax - cy)(ax + cy) = 0.$$

C'est l'*équation d'un couple de droites concourantes en l'origine des coordonnées*: les droites $ax - cy = 0$ et $ax + cy = 0$.

3) *Genre parabole.* Si $AC - B^2 = 0$, une rotation des axes de coordonnées du même angle α que dans le lemme 3.1 ramène l'équation générale à la forme

$$Ax^2 + Cy^2 + 2Ey + 2Dx + F = 0. \tag{7}$$

Comme $AC = 0$, l'un des coefficients A ou C est nul.

Supposons que $A = 0$, $C \neq 0$. Mettons l'équation (7) sous la forme

$$C\left[y^2 + \frac{2E}{C}y + \left(\frac{E}{C}\right)^2\right] + 2Dx + F - \frac{E^2}{C} = 0,$$

ou

$$C\left(y + \frac{E}{C}\right)^2 + 2Dx + F^* = 0,$$

où $F^* = F - E^2/C$. Plaçons l'origine des coordonnées au point $(0; -E/C)$, c'est-à-dire passons aux coordonnées $x' = x$, $y' = y + E/C$. On obtient l'équation

$$Cy'^2 + 2Dx' + F^* = 0.$$

Distinguons les cas suivants:

a) $D \neq 0$. Mettons l'équation sous la forme

$$Cy'^2 + 2D\left(x' + \frac{F^*}{2D}\right) = 0.$$

Plaçons l'origine des coordonnées au point $(-F^*/(2D); 0)$, c'est-à-dire passons aux coordonnées $x'' = x' + F^*/(2D)$, $y'' = y'$. On obtient l'équation

$$Cy''^2 + 2Dx'' = 0$$

ou

$$y''^2 = 2px'',$$

où $p = -D/C$. On reconnaît l'équation canonique de la parabole.

b) $D = 0$. L'équation devient

$$Cy'^2 + F^* = 0.$$

Si C et F^* sont de signes contraires, en posant $|F^*/C| = a^2$, on peut mettre cette équation sous la forme $(y' - a)(y' + a) = 0$. On reconnaît ici l'équation d'un couple de droites parallèles.

Si C et F^* sont de même signe, l'équation devient $y'^2 + a^2 = 0$. C'est l'équation d'un couple de droites parallèles imaginaires.

Enfin, si $F^* = 0$, l'équation devient $y'^2 = 0$ et définit l'axe $O'x'$. Cette équation peut être traitée comme l'équation d'un couple de droites confondues dans le cas limite où $F^* \to 0$.

Résumons tous ces résultats sous forme d'un théorème.

Théorème 3.7. *Soit donnée dans un système d'coordonnées rectangulaires l'équation générale d'une conique*

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0.$$

Il existe alors un système de coordonnées rectangulaires dans lequel cette équation prend l'une des neuf formes canoniques suivantes:

1) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ *une ellipse réelle;*

2) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1$ *une ellipse imaginaire;*

3) $a^2x^2 + c^2y^2 = 0$ *un couple de droites concourantes imaginaires;*

4) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ *une hyperbole;*

5) $a^2x^2 - c^2y^2 = 0$ *un couple de droites concourantes;*

6) $y^2 = 2px$ *une parabole;*

7) $y^2 - a^2 = 0$ *un couple de droites parallèles;*

8) $y^2 + a^2 = 0$ *un couple de droites parallèles imaginaires;*

9) $y^2 = 0$ *un couple de droites confondues.*

CHAPITRE IV

FONCTIONS D'UNE VARIABLE

On se propose d'étudier une notion très importante d'analyse mathématique: la notion de fonction. Dans ce chapitre on introduit les notions de limite et de continuité d'une fonction.

§ 1. Notion de fonction

1. Définition d'une fonction.

Définition. Soient X et Y des ensembles numériques. On appelle *fonction* l'ensemble f des couples de nombres $(x; y)$ tels que $x \in X$, $y \in Y$, le nombre x figurant dans un couple et un seul de f, le nombre y, au moins dans un couple. On dit alors qu'au nombre x *est associé* le nombre y et l'on écrit $y = f(x)$. Le nombre y s'appelle *valeur* de la fonction f au point x. La variable y s'appelle *variable dépendante*, la variable x, *variable indépendante* (ou *argument*), l'ensemble X, *ensemble de définition* (ou *d'existence*) de la fonction f, l'ensemble Y, *ensemble des valeurs de* f.

On se sert des lettres autres que f pour désigner les fonctions, par exemple: $y = y(x)$, $y = g(x)$, $y = \varphi(x)$, $y = A(x)$, $y = F(x)$, etc. Les variables dépendante et indépendante peuvent être désignées par d'autres lettres aussi. La variable dépendante est parfois appelée *fonction*.

Outre le terme de « fonction » on utilise le terme équivalent d'« application » et au lieu de $y = f(x)$ on écrit $f: x \mapsto y$ et l'on dit que l'application f associe x à y ou encore que y est l'image de x par f.

Dans les calculs la notation $y = f(x)$ est plus commode que $f: x \mapsto y$. Par exemple, la notation $f(x) = x^2$ est bien plus commode à l'usage et plus maniable dans les transformations analytiques que la notation $f: x \mapsto x^2$.

La fonction dont toutes les valeurs sont égales entre elles s'appelle *constante*. On la désigne souvent par la lettre C.

Une fonction $f(x)$ définie sur un ensemble X est dite *majorée* (resp. *minorée*) sur cet ensemble s'il existe un nombre M (resp. m) tel que pour tout $x \in X$ l'on ait $f(x) \leqslant M$ (resp. $f(x) \geqslant m$).

Une fonction $f(x)$ à la fois majorée et minorée sur X est dite *bornée* sur X.

Le fait qu'une fonction $f(x)$ est bornée peut s'écrire de la sorte: il existe un nombre $M > 0$ tel que pour tout $x \in X$ l'on a $|f(x)| \leqslant M$. Par exemple, la fonction $f(x) = \sin x$ est bornée sur la droite numérique tout entière, puisque $|\sin x| \leqslant 1$ pour tout x; la fonction $f(x) = 1/x$ n'est pas majorée sur l'intervalle $]0, 1[$, puisqu'il n'existe pas de M tel que $1/x \leqslant M$ pour tout $x \in]0, 1[$.

On nomme *courbe représentative* ou *représentation graphique* ou tout simplement *graphique* d'une fonction $f(x)$ l'ensemble des points $(x; y)$ du plan dont les coordonnées vérifient la relation $y = f(x)$, appelée *équation du graphique*.

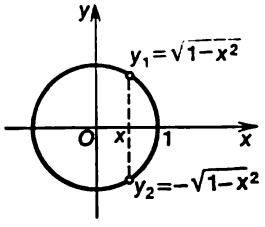


Fig. 43

Le graphique d'une fonction peut être soit une ligne « continue » (une courbe ou une droite), soit être composé de points isolés, tel le graphique de la fonction $y = n!$ (fig. 45).

Signalons qu'une courbe n'est pas nécessairement le graphique d'une fonction. Par exemple, le cercle $x^2 + y^2 = 1$ (fig. 43) n'est le graphique d'aucune fonction, puisque chaque $x \in]-1, 1[$ figure non pas dans un mais dans deux couples $(x; y)$ de cet ensemble avec différentes valeurs de y: $y_1 = \sqrt{1 - x^2}$ et $y_2 = -\sqrt{1 - x^2}$, ce qui contredit la condition d'univocité dans la définition de la fonction. Cependant la partie du cercle située dans le demi-plan inférieur est le graphique de la fonction $y = -\sqrt{1 - x^2}$, quant à la partie qui est située dans le demi-plan supérieur, elle est le graphique de la fonction $y = \sqrt{1 - x^2}$.

2. Modes de définition des fonctions. Définir une fonction f c'est indiquer le procédé par lequel à chaque valeur de x est associée la valeur correspondante $f(x)$. Il existe trois modes de définition des fonctions: le mode analytique, le mode tabulaire et le mode graphique.

Mode analytique. La dépendance des variables est exprimée par des formules indiquant les opérations à accomplir pour calculer la valeur de la fonction associée à la valeur donnée de l'argument.

Exemples.

1. La formule $y = x^2$ définit une fonction dont le domaine de définition est la droite numérique $]-\infty, +\infty[$ et l'ensemble des valeurs, la demi-droite $[0, +\infty[$ (fig. 44, *a*).

2. La formule $y = \sqrt{1 - x^2}$ définit une fonction dont le domaine de définition est le segment $[-1, 1]$ et l'ensemble des valeurs, le segment $[0, 1]$ (fig. 44, *b*).

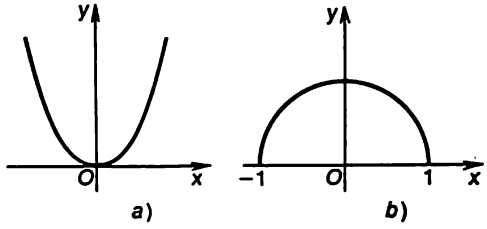


Fig. 44

3. La formule $y = n!$ associe à chaque entier naturel n le nombre $y = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$. Par exemple, si $n = 3$, $y = 3! = 6$.

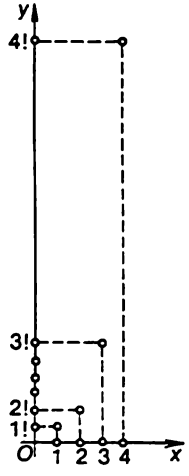


Fig. 45

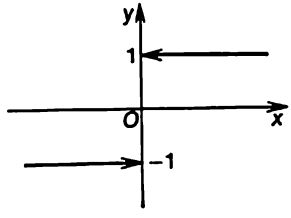


Fig. 46

La formule $y = n!$ définit une fonction dont l'ensemble de définition est $\{1, 2, 3, \ldots, n, \ldots\}$ et l'ensemble des valeurs, $\{1!, 2!, 3!, \ldots, n!, \ldots\}$ (fig. 45).

4. $y = \operatorname{sgn} x = \begin{cases} +1 \text{ si } x > 0 \\ 0 \text{ si } x = 0 \\ -1 \text{ si } x < 0. \end{cases}$

Cette fonction est donnée par plusieurs formules. Elle est définie sur toute la droite numérique $]-\infty, +\infty[$, son ensemble des valeurs est composé de trois nombres: -1, 0 et $+1$ (fig. 46).

5. La fonction de Dirichlet *)

$$y = \begin{cases} 1 \text{ si } x \text{ est rationnel} \\ 0 \text{ si } x \text{ est irrationnel.} \end{cases}$$

Cette fonction est définie sur toute la droite numérique $]-\infty, +\infty[$, son ensemble des valeurs est constitué des nombres 0 et 1.

Signalons qu'il est impossible de représenter la fonction de Dirichlet par un graphique.

Mode tabulaire. Considérons la table suivante:

x	0	0,1	0,2	3	0,6	4	0,8	1,5	2
y	−1	10	1	−2	−8	0,5	−2	5	7

A toute valeur de x associons la valeur de y située dans la même colonne et disons que la fonction y est donnée tabulairement. L'ensemble de définition de y est l'ensemble des neuf valeurs de x de la première ligne, l'ensemble des valeurs, l'ensemble des neuf nombres y de la deuxième ligne.

Le mode tabulaire ne convient que pour un nombre fini de valeurs de l'argument. Les tables sont souvent utilisées pour donner une fonction. Comme exemples classiques citons la table des fonctions trigonométriques, la table des logarithmes, etc. Indiquons aussi l'horaire des trains qui permet de localiser un train en fonction de l'heure.

Mode graphique. Le mode graphique est généralement utilisé pour les mesures en physique lorsque la correspondance entre des variables x et y est exprimée directement par un graphique. Ces graphiques sont tracés dans de nombreux cas par des autoscripteurs. Ainsi, par exemple, la pression atmosphérique à des altitudes différentes est mesurée par un instrument spécial, appelé barographe, qui trace sur un ruban une courbe représentant la pression en fonction de l'altitude.

3. Classification des fonctions. La fonction constante $f(x) = C$, $C = \text{const}$, la fonction puissance x^α (α est un nombre quelconque), la fonction exponentielle a^x ($0 < a \neq 1$), la fonction logarithme $\log_a x$ ($0 < a \neq 1$), les fonctions trigonométriques $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{cotg} x$ et leurs réciproques $\operatorname{Arc\,sin} x$, $\operatorname{Arc\,cos} x$, $\operatorname{Arc\,tg} x$ et $\operatorname{Arc\,cotg} x$ s'appellent *fonctions élémentaires.*

On obtient d'autres fonctions par un nombre fini d'opérations arithmétiques sur les fonctions élémentaires ainsi que par la

*) Dirichlet Peter Gustav Lejeune (1805-1859), mathématicien allemand.

composition de ces fonctions. Exemple

$$f(x) = |x| \; (|x| = \sqrt{x^2}); \; f(x) = \log^3 \operatorname{Arc\,tg}\, 2^{\sqrt{x}} + \\ + \sin 3x; \; f(x) = \ln |\sin 3x| - e^{\operatorname{Arc\,tg}\sqrt{x}}, \text{ etc.}$$

On a la classification suivante pour les fonctions élémentaires.

1) Une fonction de la forme

$$P(x) = a_0 x^m + a_1 x^{m-1} + \ldots + a_{m-1} x + a_m,$$

où $m \geqslant 0$ est un entier; $a_0, a_1, \ldots, a_m$ des nombres quelconques ($a_0 \neq 0$), s'appelle *fonction polynôme de degré m*. Une fonction polynôme du premier degré s'appelle aussi *fonction linéaire*.

2) Une fonction qui est le quotient de deux fonctions polynômes

$$R(x) = \frac{a_0 x^m + a_1 x^{m-1} + \ldots + a_{m-1} x + a_m}{b_0 x^n + b_1 x^{n-1} + \ldots + b_{n-1} x + b_n}$$

s'appelle *fonction rationnelle* ou *fraction rationnelle*.

3) Une fonction non rationnelle qui s'obtient par un nombre fini de compositions et par les quatre opérations arithmétiques sur des fonctions puissances à exposants aussi bien entiers que fractionnaires s'appelle *fonction irrationnelle*. Exemples: $f(x) = \sqrt{x}$, $f(x) = x + \sqrt{x}$, $f(x) = \sqrt{(5x^2 + 4x - 7)/(3x^2 - 8x + 4)} + (\sqrt[5]{x} + x)^3$, etc.

4) Une fonction qui n'est ni rationnelle, ni irrationnelle, s'appelle *transcendante*. Exemples: $f(x) = \sin x$, $f(x) = \sin x + x$, etc.

§ 2. Limite d'une fonction

1. Limite d'une fonction lorsque $x \to x_0$. Soient $f(x)$ une fonction définie sur un ensemble X et x_0 un point appartenant ou non à X. Prenons une suite de points

$$x_1, x_2, x_3, \ldots, x_n, \ldots \qquad (1)$$

de X distincts de x_0 convergeant vers x_0 *). Les valeurs de la fonction $f(x)$ en ces points forment aussi une suite numérique

$$f(x_1), f(x_2), f(x_3), \ldots, f(x_n), \ldots, \qquad (2)$$

à propos de laquelle il est légitime de se demander si elle admet une limite.

Définition 1. On dit qu'un nombre A est la *limite* d'une fonction $f(x)$ en $x = x_0$ (ou lorsque $x \to x_0$) si pour toute suite $\{x_n\}$ convergeant vers x_0, la suite $\{f(x_n)\}$ converge vers A.

Ceci se note: $\lim\limits_{x \to x_0} f(x) = A$.

*) On admet qu'une telle suite existe.

La fonction $f(x)$ ne peut admettre qu'une seule limite en x_0. Ceci résulte du fait que la suite $\{f(x_n)\}$ ne possède qu'une seule limite.

Exemples.

1. La fonction $f(x) = C$ admet une limite en chaque point x_0 de la droite numérique. En effet, si (1) est une suite quelconque

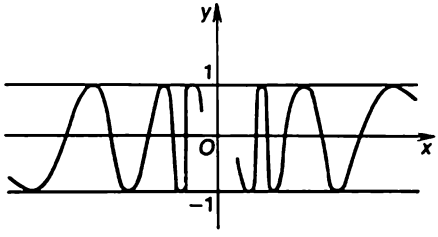


Fig. 47

convergeant vers x_0, la suite (2) est de la forme $C, C, C, \ldots$ $\ldots, C, \ldots$, c'est-à-dire que $f(x_n) = C$. D'où l'on déduit que $f(x_n) \to C$ lorsque $n \to \infty$ ou $\lim\limits_{x \to x_0} f(x) = C$.

2. La fonction $f(x) = x$ admet en tout point x_0 de la droite numérique une limite égale à x_0. Dans ce cas les suites (1) et (2) sont identiques, c'est-à-dire que $f(x_n) = x_n$. Donc si $x_n \to x_0$, il en est de même de $f(x_n)$, ce qu'on écrit encore

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} x = f(x_0) = x_0.$$

3. La fonction $f(x) = \sin(1/x)$ (fig. 47) qui est définie pour tout $x \neq 0$ ne possède pas de limite pour $x = 0$. En effet, prenons deux suites de valeurs de l'argument x:

$\frac{1}{\pi}, \frac{1}{2\pi}, \frac{1}{3\pi}, \ldots, \frac{1}{n\pi}, \ldots$ et $\frac{2}{\pi}, \frac{2}{5\pi}, \frac{2}{9\pi}, \ldots, \frac{2}{(4n-3)\pi}, \ldots$

convergeant vers 0. Les suites correspondantes des valeurs de la fonction sont $f\left(\frac{1}{\pi}\right), f\left(\frac{1}{2\pi}\right), f\left(\frac{1}{3\pi}\right), \ldots, f\left(\frac{1}{n\pi}\right), \ldots$ et $f\left(\frac{2}{\pi}\right), f\left(\frac{2}{5\pi}\right), f\left(\frac{2}{9\pi}\right), \ldots, f\left(\frac{2}{(4n-3)\pi}\right), \ldots$ Puisque $f\left(\frac{1}{n\pi}\right) = \sin n\pi = 0$ et $f\left(\frac{2}{(4n-3)\pi}\right) = \sin\frac{(4n-3)\pi}{2} = 1$, il vient

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} \sin n\pi = 0 \text{ et } \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} \sin\frac{(4n-3)\pi}{2} = 1.$$

Donc la fonction $f(x)$ tend vers des limites différentes pour des suites $\{x_n\}$ tendant toutes deux vers 0. Ce qui exprime par définition de la limite d'une fonction que $\lim\limits_{x \to 0} f(x)$ n'existe pas.

4. La fonction $f(x) = \frac{x^2+x-1}{x-1}$ admet au point $x = 0$ une

limite égale à 1. En effet, prenons une suite quelconque de valeurs de l'argument x convergeant vers 0: $\lim\limits_{n\to\infty} x_n = 0$, $x_n \neq 0$. En vertu des théorèmes 2.7 à 2.9, on a alors

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} \frac{x_n^2+x_n-1}{x_n-1} = \frac{(\lim\limits_{n\to\infty} x_n)^2+\lim\limits_{n\to\infty} x_n-1}{\lim\limits_{n\to\infty} x_n-1} = 1.$$

Donc $\lim\limits_{n\to\infty} f(x_n)$ existe et comme elle ne dépend pas du choix de la suite $\{x_n\}$, on conclut aux termes de la définition de la limite d'une fonction que $\lim\limits_{x\to 0} f(x) = 1$.

5. La fonction de Dirichlet qui est égale à 1 pour les points rationnels et à 0 pour les points irrationnels ne possède de limite en aucun point x_0 de la droite numérique. En effet, la suite $\{f(x_n)\}$ tend vers 1 ou 0 selon que la suite $\{x_n\}$ convergeant vers x_0 est composée de points rationnels ou irrationnels.

Il existe une autre définition de la limite d'une fonction.

Définition 2. On dit qu'un nombre A est la *limite* d'une fonction $f(x)$ en $x = x_0$ si pour tout nombre $\varepsilon > 0$ il existe un nombre $\delta > 0$ tel que pour tout $x \in X$, $x \neq x_0$, satisfaisant l'inégalité $|x - x_0| < \delta$ on ait $|f(x) - A| < \varepsilon$.

En signes logiques ceci s'écrit:

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in X,\ x \neq x_0,\ |x - x_0| < \delta: \quad |f(x) - A| < \varepsilon.$$

Signalons que les inégalités $x \neq x_0$, $|x - x_0| < \delta$ sont équivalentes à la double inégalité $0 < |x - x_0| < \delta$.

La première définition est basée sur la notion de limite d'une suite numérique, aussi l'appelle-t-on souvent définition dans le « langage des suites ». La deuxième définition s'appelle définition dans le « langage de ε, δ ».

Théorème 4.1. *Les définitions* 1 *et* 2 *de la limite d'une fonction sont équivalentes.*

Démonstration. 1) Soit A la limite d'une fonction $f(x)$ en x_0 aux termes de la première définition. Montrons que A est la limite aux termes de la deuxième définition. Supposons par absurde que A n'est pas la limite de $f(x)$ aux termes de la deuxième définition. Cela signifie que ce n'est pas pour tout $\varepsilon > 0$ qu'on peut exhiber un $\delta > 0$ tel que la double inégalité $0 < |x - x_0| < \delta$ entraîne $|f(x) - A| < \varepsilon$, c'est-à-dire qu'il existe un $\varepsilon = \varepsilon_0 > 0$ pour lequel on peut, quel que soit $\delta > 0$, trouver au moins un point $x \neq x_0$ tel que $|x - x_0| < \delta$, mais $|f(x) - A| \geqslant \varepsilon_0$. Pour δ prenons successivement les nombres:

$$1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$$

Alors:
pour $\delta = 1$ il existe dans X un $x_1 \neq x_0$ tel que $|x_1 - x_0| < 1$ et $|f(x_1) - A| \geqslant \varepsilon_0$;
pour $\delta = 1/2$ il existe dans X un $x_2 \neq x_0$ tel que $|x_2 - x_0| < 1/2$ et $|f(x_2) - A| \geqslant \varepsilon_0$;
pour $\delta = 1/3$ il existe dans X un $x_3 \neq x_0$ tel que $|x_3 - x_0| < 1/3$ et $|f(x_3) - A| \geqslant \varepsilon_0$;
. . . . . .
pour $\delta = 1/n$ il existe dans X un $x_n \neq x_0$ tel que $|x_n - x_0| < 1/n$ et $|f(x_n) - A| \geqslant \varepsilon_0$.
. . . . . .

On obtient en définitive une suite de points

$$x_1, x_2, x_3, \ldots, x_n, \ldots,$$

distincts de x_0 convergeant vers x_0, puisque $|x_n - x_0| < 1/n \to 0$ lorsque $n \to \infty$. Donc aux termes de la première définition d'une fonction, la suite $\{f(x_n)\}$ tend vers A. Donc pour ε_0 on peut exhiber un entier N tel que pour tout $n > N$ l'on ait $|f(x_n) - A| < \varepsilon_0$. Or ceci est impossible, puisque $|f(x_n) - A| \geqslant \varepsilon_0$ pour tout x_n. Cette contradiction prouve que A est la limite de $f(x)$ en x_0 aux termes de la deuxième définition.

2) Supposons maintenant que A est la limite de $f(x)$ en x_0 aux termes de la deuxième définition. Cela signifie que pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que la double inégalité $0 < |x - x_0| < \delta$ entraîne $|f(x) - A| < \varepsilon$. Montrons que A est la limite de $f(x)$ aux termes de la première définition. Prenons une suite quelconque de points

$$x_1, x_2, x_3, \ldots, x_n, \ldots,$$

convergeant vers $x_0 \neq x_n$. Pour la valeur de $\delta > 0$ associée à ε dans le cadre de la deuxième définition, il existe un entier N tel que pour tout $n > N$ l'on a $0 < |x_n - x_0| < \delta$. Mais parallèlement, aux termes de la deuxième définition, on aura $|f(x_n) - A| < \varepsilon$. Or, comme ε est arbitraire, cela exprime que $f(x_n) \to A$ pour toute suite $\{x_n\}$ convergeant vers $x_0 \neq x_n$, autrement dit le nombre A est la limite de $f(x)$ en x_0 aux termes de la première définition. ■

On a ainsi établi l'équivalence des deux définitions de la limite d'une fonction et l'on peut se servir indifféremment de l'une ou de l'autre selon les avantages procurés dans la résolution des problèmes.

Signalons que la définition 1 est la définition de la limite d'une fonction au sens de Heine *), la définition 2, au sens de Cauchy **).

*) Heine Heinrich Eduard (1821-1881), mathématicien allemand.
**) Cauchy Augustin Louis (1789-1857), mathématicien français.

2. Limite d'une fonction lorsque x tend vers x_0 à gauche et à droite. On se servira ultérieurement des notions de limites à gauche et à droite d'une fonction qui se définissent comme suit.

Définition 3. On dit qu'un nombre A est la *limite à droite* (resp. *à gauche*) d'une fonction $f(x)$ en x_0 si pour toute suite $\{x_n\}$ convergeant vers x_0 par valeurs décroissantes (resp. croissantes), la suite $\{f(x_n)\}$ converge vers A.

Ceci se note: $\lim\limits_{x \to x_0+} f(x) = A$ (resp. $\lim\limits_{x \to x_0-} f(x) = A$).

Voyons à titre d'exemple la fonction $f(x) = \operatorname{sgn} x$. Cette fonction admet une limite à droite et une à gauche au point $x = 0$: $\lim\limits_{x \to 0+} \operatorname{sgn} x = 1$, $\lim\limits_{x \to 0-} \operatorname{sgn} x = -1$. En effet, si $\{x_n\}$ est une suite quelconque convergeant vers 0 par valeurs positives, alors $\operatorname{sgn} x_n = 1$ et $\lim\limits_{n \to \infty} \operatorname{sgn} x_n = 1$. Donc, $\lim\limits_{x \to 0+} \operatorname{sgn} x = 1$. On établit de façon analogue que $\lim\limits_{x \to 0-} \operatorname{sgn} x = -1$.

On peut donner une définition équivalente des limites à droite et à gauche en termes de ε et δ: on dit qu'un nombre A est la limite à droite (resp. à gauche) d'une fonction $f(x)$ en x_0 si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout x vérifiant la double inégalité $x_0 < x < x_0 + \delta$ (resp. $x_0 - \delta < x < x_0$) l'on ait $|f(x) - A| < \varepsilon$.

En signes logiques cela devient:

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \forall x, \quad x_0 < x < x_0 + \delta \ (x_0 - \delta < x < x_0): |f(x) - A| < \varepsilon.$$

Le théorème suivant établit un lien entre les limites à droite et à gauche et la limite d'une fonction.

Théorème 4.2. *Une condition nécessaire et suffisante pour qu'une fonction $f(x)$ admette une limite en un point x_0 est l'existence et l'égalité des limites à droite et à gauche en ce point.*

Démonstration. Soit $\lim\limits_{x \to x_0-} f(x) = \lim\limits_{x \to x_0+} f(x) = A$. Par définition de la limite à gauche et de la limite à droite d'une fonction, pour tout $\varepsilon > 0$ il existe des nombres $\delta_1 > 0$ et $\delta_2 > 0$ tels que pour tout x vérifiant la double inégalité $x_0 - \delta_1 < x < x_0$ et tout x vérifiant la double inégalité $x_0 < x < x_0 + \delta_2$, l'on a $|f(x) - A| < \varepsilon$. Prenons $\delta = \min\{\delta_1, \delta_2\}$. Pour tout x satisfaisant l'inégalité $|x - x_0| < \delta$, $x \neq x_0$, on aura alors $|f(x) - A| < \varepsilon$. Or ceci exprime d'après la définition 2, que $\lim\limits_{x \to x_0} f(x) = A$.

Réciproquement, supposons que $\lim\limits_{x \to x_0} f(x) = A$. En vertu de la définition de la limite d'une fonction en x_0, pour tout $\varepsilon > 0$ il existe un nombre $\delta > 0$ tel que pour tout x vérifiant

l'inégalité $|x - x_0| < \delta$, $x \neq x_0$, l'on a $|f(x) - A| < \varepsilon$. Donc $|f(x) - A| < \varepsilon$ aussi bien pour $x_0 - \delta < x < x_0$ que pour $x_0 < x < x_0 + \delta$. Ce qui aux termes de la définition des limites à gauche et à droite exprime que $\lim\limits_{x \to x_0 -} f(x) = \lim\limits_{x \to x_0 +} f(x) = A$. ■

3. Limite d'une fonction à l'infini. Outre les notions de limite d'une fonction lorsque $x \to x_0$ et de limites à droite et à gauche, il existe la notion de limite d'une fonction lorsque l'argument tend vers l'infini.

Définition 4. On dit qu'un nombre A est la *limite d'une fonction* $f(x)$ lorsque $x \to \infty$ si pour toute suite $\{x_n\}$ infiniment grande, la suite $\{f(x_n)\}$ converge vers A.

On note ceci: $\lim\limits_{x \to \infty} f(x) = A$.

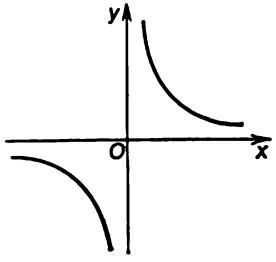


Fig. 48

Définition 5. On dit qu'un nombre A est la *limite d'une fonction* $f(x)$ lorsque $x \to +\infty$ (resp. $-\infty$) si pour toute suite $\{x_n\}$ infiniment grande, la suite $\{f(x_n)\}$ converge vers A.

On note ceci: $\lim\limits_{x \to +\infty} f(x) = A$ (resp. $\lim\limits_{x \to -\infty} f(x) = A$).

Exemple. Soit $f(x) = \frac{1}{x}$. Cette fonction tend vers 0 lorsque x tend vers l'infini. En effet, si $\{x_n\}$ est une suite infiniment grande, la suite $\{f(x_n)\} = \left\{\frac{1}{x_n}\right\}$ est infiniment petite d'après le théorème 2.1, donc elle tend vers 0: $\lim\limits_{x \to \infty} \frac{1}{x} = 0$ (fig. 48).

Les définitions 4 et 5 sont données dans le langage des suites. On peut produire des définitions analogues dans le langage de ε et δ et les écrire avec des signes logiques. Nous laissons ceci au soin du lecteur. A titre d'exemple formulons la définition de la limite d'une fonction lorsque $x \to +\infty$.

Définition 6. On dit qu'un nombre A est *la limite d'une fonction* $f(x)$ *lorsque* $x \to +\infty$ si pour tout $\varepsilon > 0$ il existe un nombre δ tel que pour tout $x \in X$ vérifiant l'inégalité $x > \delta$ l'on ait $|f(x) - A| < \varepsilon$.

§ 3. Théorèmes relatifs aux limites des fonctions

La définition de la limite d'une fonction dans le langage des suites permet de transposer aux fonctions les théorèmes prouvés plus haut pour les limites des suites. Montrons ceci sur l'exemple de deux théorèmes.

Théorème 4.3. *Soient $f(x)$ et $g(x)$ des fonctions possédant en x_0 des limites respectivement égales à B et C. Alors les fonctions $f(x) \pm g(x)$, $f(x)\,g(x)$ et $\frac{f(x)}{g(x)}$ $(C \neq 0)$ possèdent aussi en x_0 des limites respectivement égales à $B \pm C$, BC et $\frac{B}{C}$.*

Démonstration. Soit $\{x_n\}$ $(x_n \neq x_0)$ une suite quelconque de valeurs de x convergeant vers x_0. Les suites $\{f(x_n)\}$ et $\{g(x_n)\}$ convergent vers B et C. D'après les théorèmes 2.7 à 2.9 les suites $\{f(x_n) \pm g(x_n)\}$, $\{f(x_n)\,g(x_n)\}$ et $\{f(x_n)/g(x_n)\}$ $(C \neq 0)$ admettent des limites respectivement égales à $B \pm C$, BC et B/C. Aux termes de la définition 1 de la limite d'une fonction cela signifie que

$$\lim_{x \to x_0} [f(x) \pm g(x)] = B \pm C, \qquad \lim_{x \to x_0} [f(x)\,g(x)] = BC,$$

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{B}{C}. \quad \blacksquare$$

Théorème 4.4. *Supposons que des fonctions $f(x)$, $g(x)$ et $h(x)$ sont définies dans un voisinage de x_0 sauf éventuellement en x_0 et que $\lim\limits_{x \to x_0} f(x) = \lim\limits_{x \to x_0} h(x) = A$. Si par ailleurs $f(x) \leqslant g(x) \leqslant h(x)$, alors*

$$\lim_{x \to x_0} g(x) = A.$$

Démonstration. Soit $\{x_n\}$ $(x_n \neq x_0)$ une suite quelconque convergeant vers x_0. Les suites $\{f(x_n)\}$ et $\{h(x_n)\}$ tendent vers A lorsque $x_n \to x_0$, c'est-à-dire que $f(x_n) \to A$ et $h(x_n) \to A$ pour $n \to \infty$. Par hypothèse,

$$f(x_n) \leqslant g(x_n) \leqslant h(x_n).$$

D'où $g(x_n) \to A$ en vertu du théorème 2.11. Aux termes de la définition 1 de la limite d'une fonction, cela exprime que

$$\lim_{x \to x_0} g(x) = A. \quad \blacksquare$$

Remarque. Les théorèmes 4.3 et 4.4 sont valables aussi dans le cas où $x_0 = \pm \infty$.

§ 4. Deux limites remarquables

1. $\lim\limits_{x\to 0}\dfrac{\sin x}{x}$. Montrons que

$$\lim_{x\to 0}\frac{\sin x}{x}=1.$$

Considérons l'arc de cercle de rayon $R=1$ intercepté par un angle au centre de mesure x en radians $\left(0<x<\frac{\pi}{2}\right)$ (fig. 49). Alors

$$OA=1,\ \sin x=MK,\ \operatorname{tg} x=AT. \qquad (1)$$

Il est évident que l'aire du triangle OAM est strictement inférieure à celle du secteur OAM, laquelle est strictement inférieure à celle du triangle OAT, soit $\frac{OA\cdot MK}{2}<\frac{OA\cdot \overset{\smile}{AM}}{2}<\frac{OA\cdot AT}{2}$. La relation (1) nous permet d'écrire cette double inégalité sous la forme $\frac{\sin x}{2}<\frac{x}{2}<\frac{\operatorname{tg} x}{2}$ d'où

$$\sin x<x<\operatorname{tg} x. \qquad (2)$$

En divisant par $\sin x$ on trouve

$$1>\frac{\sin x}{x}>\cos x,$$

d'où

$$0<1-\frac{\sin x}{x}<1-\cos x.$$

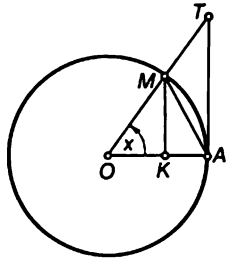


Fig. 49

Vu que $\sin\frac{x}{2}<1$, on a $\sin^2\frac{x}{2}<\sin\frac{x}{2}$. Donc en tenant compte de la première inégalité (2), pour tout x compris entre 0 et $\frac{\pi}{2}$ on a

$$1-\cos x=2\sin^2\frac{x}{2}<2\sin\frac{x}{2}<2\,\frac{x}{2}=x.$$

Ainsi

$$0<1-\frac{\sin x}{x}<x$$

pour $0<x<\frac{\pi}{2}$.

Prenons un $\varepsilon>0$ arbitraire et posons $\delta=\min\left\{\varepsilon,\frac{\pi}{2}\right\}$. Pour tout x tel que $0<x<\delta$, on aura $x<\varepsilon$ et par suite

$$0<1-\frac{\sin x}{x}<\varepsilon \text{ d'où } \left|1-\frac{\sin x}{x}\right|<\varepsilon.$$

Ce qui exprime que 1 est la limite à droite de la fonction $\frac{\sin x}{x}$ au point $x=0$: $\lim\limits_{x\to 0+}\frac{\sin x}{x}=1$. Remarquons maintenant que la fonction $f(x)=\frac{\sin x}{x}$ est paire, puisque $f(-x)=\frac{\sin(-x)}{-x}=$ $=\frac{\sin x}{x}=f(x)$. Donc la limite à gauche de la fonction $\frac{\sin x}{x}$ est aussi égale à 1 en $x=0$. Il en résulte en vertu du théorème 4.2 que $\lim\limits_{x\to 0}\frac{\sin x}{x}=1$. ■

Remarque. En utilisant les inégalités $\sin x<x$ et $1-\cos x<x$ lorsque $0<x<\frac{\pi}{2}$ obtenues ci-dessus, on démontre sans peine que $\lim\limits_{x\to 0}\cos x=1$ et $\lim\limits_{x\to 0}\sin x=0$. On laisse ceci au soin du lecteur.

Cette limite permet le calcul de nombreuses autres limites.

Exemple 1. Trouver $\lim\limits_{x\to 0}\frac{1-\cos x}{x}$.

Solution. Le théorème 4.3 ne passe pas, puisque le dénominateur tend vers 0 lorsque $x\to 0$. Transformons cette fraction:

$$\lim_{x\to 0}\frac{1-\cos x}{x}=\lim_{x\to 0}\frac{2\sin^2(x/2)}{x}=\lim_{x\to 0}\frac{\sin(x/2)}{x/2}\sin(x/2)=$$

$$=\lim_{x\to 0}\frac{\sin(x/2)}{x/2}\lim_{x\to 0}\sin(x/2)=1\cdot 0=0.$$

Exemple 2. Trouver $\lim\limits_{x\to 0}\frac{\operatorname{tg} x}{x}$.

Solution. On a

$$\lim_{x\to 0}\frac{\operatorname{tg} x}{x}=\lim_{x\to 0}\frac{\sin x}{x}\frac{1}{\cos x}=\lim_{x\to 0}\frac{\sin x}{x}\lim_{x\to 0}\frac{1}{\cos x}=1\cdot\frac{1}{1}=1.$$

Exemple 3. Trouver $\lim\limits_{x\to 0}\frac{5x}{\sin 4x}$.

Solution. On a

$$\lim_{x\to 0}\frac{5x}{\sin 4x}=\lim_{x\to 0}\frac{5/4}{(\sin 4x)/(4x)}=\frac{5/4}{\lim\limits_{x\to\infty}(\sin 4x)/(4x)}=\frac{5/4}{1}=1{,}25.$$

2. $\boxed{\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^x.}$ Montrons que $\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e$. On sait que $\lim\limits_{n\to\infty}(1+1/n)^n=e$ (cf. chap. 2, § 3, n° 2). Supposons que

$x > 1$. Posons $n = [x]$; alors $x = n + \alpha$, où n est un entier naturel et α, un nombre vérifiant la condition $0 \leqslant \alpha < 1$. Puisque $n \leqslant x < n+1$, on a $\frac{1}{n+1} < \frac{1}{x} \leqslant \frac{1}{n}$ et

$$\left(1 + \frac{1}{n+1}\right)^n < \left(1 + \frac{1}{x}\right)^x < \left(1 + \frac{1}{n}\right)^{n+1}.$$

Lorsque $x \to +\infty$ $(n \to \infty)$

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^{n+1} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = e \cdot 1 = e$$

et

$$\lim_{n\to\infty} \left(1 + \frac{1}{n+1}\right)^n = \frac{\lim\limits_{n\to\infty} \left(1 + \frac{1}{n+1}\right)^{n+1}}{\lim\limits_{n\to\infty} \left(1 + \frac{1}{n+1}\right)} = \frac{e}{1} = e.$$

Il en résulte d'après le théorème 4.4

$$\lim_{x\to+\infty} \left(1 + \frac{1}{x}\right)^x = e.$$

Supposons maintenant que $x < -1$; posons $x = -y$. Alors

$$\lim_{x\to-\infty} \left(1 + \frac{1}{x}\right)^x = \lim_{y\to+\infty} \left(1 - \frac{1}{y}\right)^{-y} = \lim_{y\to+\infty} \left(1 + \frac{1}{y-1}\right)^y =$$

$$= \lim_{y\to+\infty} \left(1 + \frac{1}{y-1}\right)^{y-1} \lim_{y\to+\infty} \left(1 + \frac{1}{y-1}\right) = e \cdot 1 = e$$

lorsque $x \to -\infty$.

En combinant les deux cas, on obtient en définitive

$$\lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x = e. \blacksquare$$

Cette deuxième limite remarquable est aussi largement appliquée dans le calcul des limites.

Exemple 4. Calculer $\lim\limits_{x\to 0} (1 + x)^{1/x}$.

▶ Solution. Faisons le changement $1/x = \alpha$. Il est alors évident que $\alpha \to \infty$ lorsque $x \to 0$. Donc

$$\lim_{x\to 0} (1+x)^{1/x} = \lim_{\alpha\to\infty} (1 + 1/\alpha)^\alpha = e.$$

Exemple 5. Calculer $\lim\limits_{x\to\infty} (1 + 3/x)^x$.

Solution. Posons $x = 3t$. On voit que $x \to \infty$ avec t. Donc

$$\lim_{x\to\infty}(1+3/x)^x = \lim_{t\to\infty}(1+1/t)^{3t} =$$

$$= \lim_{t\to\infty}[(1+1/t)^t(1+1/t)^t(1+1/t)^t] =$$

$$= \lim_{t\to\infty}(1+1/t)^t \lim_{t\to\infty}(1+1/t)^t \lim_{t\to\infty}(1+1/t)^t = e\cdot e\cdot e = e^3.$$

Exemple 6. Calculer $\lim\limits_{x\to 0}\dfrac{\log_a(1+x)}{x}$.

Solution. Transformons cette fraction comme suit:

$$\lim_{x\to 0}\frac{\log_a(1+x)}{x} = \lim_{x\to 0}\left[\frac{1}{x}\log_a(1+x)\right] =$$

$$= \lim_{x\to 0}\log_a(1+x)^{1/x} = \log_a\left[\lim_{x\to 0}(1+x)^{1/x}\right].$$

Or $\lim\limits_{x\to 0}(1+x)^{1/x} = e$ (cf. exemple 4). Donc $\lim\limits_{x\to 0}\dfrac{\log_a(1+x)}{x} =$ $= \log_a e$. En particulier, $\lim\limits_{x\to 0}\dfrac{\ln(1+x)}{x} = 1$ lorsque $a = e$.

§ 5. Infiniment petits et infiniment grands

1. Infiniment petits.

Définition 1. On dit qu'une fonction $f(x)$ est une *fonction infiniment petite* (ou, simplement, est un infiniment petit) en un point $x = x_0$ (ou lorsque $x \to x_0$) si $\lim\limits_{x\to x_0} f(x) = 0$.

On définit de façon analogue les infiniment petits lorsque $x \to \pm\infty$, $x \to x_0 -$ et $x \to x_0 +$.

On peut donner une définition équivalente en termes de ε et δ: on dit qu'une fonction $f(x)$ est un *infiniment petit* en un point $x = x_0$ si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout $x \in X$, $x \neq x_0$, vérifiant l'inégalité $|x - x_0| < \delta$ l'on ait $|f(x)| < \varepsilon$; en signes logiques cela devient:

$$\forall\varepsilon > 0,\ \exists\delta > 0,\ \forall x \in X,\ x \neq x_0,\ |x - x_0| < \delta;\ |f(x)| < \varepsilon;$$

et en termes de suites: on dit qu'une fonction $f(x)$ est un *infiniment petit* en un point $x = x_0$ si pour toute suite $\{x_n\}$, $x_n \neq x_0$, convergeant vers x_0, la suite $\{f(x_n)\}$ est un infiniment petit.

On a le théorème suivant.

Théorème 4.5. *Pour que* $\lim\limits_{x\to x_0} f(x) = A$, *il est nécessaire et suffisant que la fonction*

$$\alpha(x) = f(x) - A$$

soit un infiniment petit pour $x \to x_0$.

Démonstration. *Nécessité.* Supposons que $\lim\limits_{x \to x_0} f(x) = A$. Considérons la différence $f(x) - A = \alpha(x)$ et montrons que $\alpha(x)$ est un infiniment petit lorsque $x \to x_0$. En effet, chacune des fonctions $f(x)$ et A tend vers A lorsque $x \to x_0$. Donc en vertu du théorème 4.3

$$\lim_{x \to x_0} \alpha(x) = \lim_{x \to x_0} [f(x) - A] = \lim_{x \to x_0} f(x) - \lim_{x \to x_0} A = A - A = 0.$$

Suffisance. Supposons que $f(x) - A = \alpha(x)$, où $\alpha(x)$ est un infiniment petit lorsque $x \to x_0$. Montrons que $\lim\limits_{x \to x_0} f(x) = A$. Puisque $f(x) = A + \alpha(x)$, il vient

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} [A + \alpha(x)] = \lim_{x \to x_0} A + \lim_{x \to x_0} \alpha(x) = A + 0 = A. \blacksquare$$

Le théorème 4.5 nous donne la possibilité de représenter les fonctions possédant une limite égale à A en $x = x_0$ sous la forme spéciale:

$$f(x) = A + \alpha(x), \quad \text{où} \quad \lim_{x \to x_0} \alpha(x) = 0.$$

Ceci étant, on dit que la fonction $f(x)$ diffère de A d'un infiniment petit au voisinage du point x_0.

Les fonctions infiniment petites sont douées des mêmes propriétés que les suites infiniment petites. On a le théorème suivant.

Théorème 4.6. *La somme algébrique et le produit d'un nombre fini d'infiniment petits ainsi que le produit d'un infiniment petit par une fonction bornée sont des infiniment petits.*

Ce théorème résulte directement de la première définition de la limite d'une fonction et des théorèmes 2.2 à 2.4.

Tout ce qui a été dit au sujet des infiniment petits lorsque $x \to x_0$ reste en vigueur lorsque $x \to \pm \infty$, $x \to x_0 -$, $x \to x_0 +$.

2. Infiniment grands.

Définition 2. On dit qu'une fonction $f(x)$ est une *fonction infiniment grande* (ou, simplement, un infiniment grand) en un point $x = x_0$ (ou lorsque $x \to x_0$) si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout $x \in X$, $x \neq x_0$, vérifiant la condition $|x - x_0| < \delta$, l'on ait $|f(x)| > \varepsilon$.

Dans ce cas on écrit $\lim\limits_{x \to x_0} f(x) = \infty$ et on dit que la fonction $f(x)$ tend vers l'infini lorsque x tend vers x_0 ou encore qu'elle possède une limite infinie en $x = x_0$.

Si $f(x) > \varepsilon$ (resp. $f(x) < -\varepsilon$), on écrit $\lim\limits_{x \to x_0} f(x) = +\infty$ (resp. $\lim\limits_{x \to x_0} f(x) = -\infty$) et on dit que la fonction $f(x)$ admet en x_0 une limite infinie égale à $+\infty$ (resp. $-\infty$).

En signes logiques la définition 2 s'écrit:

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in X,\ x \neq x_0,\ |x - x_0| < \delta: \quad |f(x)| > \varepsilon.$$

Les limites infinies à droite et à gauche

$$\lim_{x \to x_0+} f(x) = +\infty, \quad \lim_{x \to x_0+} f(x) = -\infty,$$

$$\lim_{x \to x_0-} f(x) = +\infty, \quad \lim_{x \to x_0-} f(x) = -\infty$$

se définissent par analogie avec les limites homologues finies.

Ainsi, on écrit $\lim\limits_{x \to x_0+} f(x) = +\infty$ si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout $x \in X$ vérifiant la double inégalité $x_0 < x < x_0 + \delta$, on a $f(x) > \varepsilon$. En écriture logique:

$$\forall \varepsilon > 0,\ \exists \delta > 0, \quad \forall x \in X,\ x_0 < x < x_0 + \delta:\ f(x) > \varepsilon.$$

En termes de suites cette définition se formule comme suit: $\lim\limits_{x \to x_0+} f(x) = +\infty$ si pour toute suite $\{x_n\}$ convergeant vers x_0 par valeurs décroissantes, la suite $\{f(x_n)\}$ est un infiniment grand positif.

La définition rigoureuse des autres limites est laissée au soin du lecteur.

On définit de façon analogue les infiniment grands lorsque $x \to \infty$, $x \to \pm\infty$. Par exemple, on dit qu'une fonction $f(x)$ est infiniment grande lorsque $x \to \infty$ si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout $x \in X$ vérifiant l'inégalité $|x| > \delta$, on ait $|f(x)| > \varepsilon$. On note ceci: $\lim\limits_{x \to \infty} f(x) = \infty$.

En signes logiques cette définition s'énonce:

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in X,\ |x| > \delta:\ |f(x)| > \varepsilon.$$

Si $f(x) > \varepsilon$ (resp. $f(x) < -\varepsilon$), on écrit $\lim\limits_{x \to \infty} f(x) = +\infty$ (resp. $\lim\limits_{x \to \infty} f(x) = -\infty$).

On laisse au lecteur le soin de formuler la définition d'un infiniment grand lorsque $x \to \pm\infty$.

Montrons en conclusion qu'entre les infiniment grands et les infiniment petits nous avons la même relation qu'entre les suites infiniment grandes et les suites infiniment petites, autrement dit l'inverse d'un infiniment petit est un infiniment grand et vice versa.

En effet, supposons que $\lim\limits_{x \to x_0} f(x) = 0$ et $f(x) \neq 0$ pour $x \neq x_0$. Montrons que $\lim\limits_{x \to x_0} \dfrac{1}{f(x)} = \infty$. Donnons-nous arbitrairement $\varepsilon > 0$. Puisque $f(x)$ est un infiniment petit en x_0, pour le nombre $1/\varepsilon$ il existe un $\delta > 0$ tel que pour tous les $x \in X$

vérifiant la double inégalité $0 < |x - x_0| < \delta$ on a $|f(x)| < 1/\varepsilon$. Mais pour ces x on a alors $|1/f(x)| > \varepsilon$, c'est-à-dire que $1/f(x)$ est un infiniment grand en $x = x_0$. C.Q.F.D.

La réciproque est laissée au soin du lecteur.

§ 6. Comparaison des infiniment petits et des infiniment grands

Comme déjà prouvé, la somme, la différence et le produit d'infiniment petits sont des infiniment petits. Ceci n'est généralement pas vrai du quotient: la division d'un infiniment petit par un autre nous conduit à des résultats différents. Par exemple, si $\alpha(x) = x$, $\beta(x) = 2x$, alors

$$\lim_{x \to 0} \frac{\alpha(x)}{\beta(x)} = \lim_{x \to 0} \frac{x}{2x} = \frac{1}{2}.$$

Si $\alpha(x) = x$, $\beta(x) = x^2$ alors

$$\lim_{x \to 0} \frac{\alpha(x)}{\beta(x)} = \lim_{x \to 0} \frac{1}{x} = \infty, \quad \lim_{x \to 0} \frac{\beta(x)}{\alpha(x)} = \lim_{x \to 0} x = 0.$$

Considérons les règles de comparaison des infiniment petits. Soient $\alpha(x)$ et $\beta(x)$ des infiniment petits en $x \to x_0$. Alors

1) si $\lim\limits_{x \to x_0} \frac{\alpha(x)}{\beta(x)} = 0$, $\alpha(x)$ est un *infiniment petit d'ordre supérieur à* $\beta(x)$ (on dit aussi que $\alpha(x)$ a un ordre de petitesse supérieur à $\beta(x)$ lorsque $x \to x_0$);

2) si $\lim\limits_{x \to x_0} \frac{\alpha(x)}{\beta(x)} = A \neq 0$, $\alpha(x)$ et $\beta(x)$ sont des *infiniment petits du même ordre*;

3) si $\lim\limits_{x \to x_0} \frac{\alpha(x)}{\beta(x)} = 1$, $\alpha(x)$ et $\beta(x)$ sont des *infiniment petits équivalents*. L'équivalence se note: $\alpha(x) \sim \beta(x)$.

Dans certains cas il ne suffit pas de savoir qu'un infiniment petit est d'ordre supérieur à un autre, il faut encore connaître l'ordre. C'est pourquoi on introduit la règle suivante:

4) si $\lim\limits_{x \to x_0} \frac{\alpha(x)}{\beta^n(x)} = A \neq 0$, $\alpha(x)$ est un *infiniment petit d'ordre n* (par rapport à $\beta(x)$).

Les règles de comparaison des infiniment petits lorsque $x \to \infty$, $x \to \pm\infty$ et $x \to x_0$ à droite et à gauche sont identiques.

Exemples.

1. Les fonctions $\sin x$ et x sont des infiniment petits équivalents lorsque $x \to 0$, puisque $\lim\limits_{x \to 0} \frac{\sin x}{x} = 1$.

2. Les fonctions $\sin 3x$ et $\sin x$ sont des infiniment petits du même ordre lorsque $x \to 0$, puisque

$$\lim_{x\to 0} \frac{\sin 3x}{\sin x} = \lim_{x\to 0} \frac{(3\sin 3x)/(3x)}{(\sin x)/x} = 3 \lim_{x\to 0} \frac{\sin 3x}{3x} \lim_{x\to 0} \frac{x}{\sin x} = 3.$$

3. La fonction $\alpha(x) = 1 - \cos x$ est un infiniment petit d'ordre deux par rapport à x lorsque $x \to 0$, puisque

$$\lim_{x\to 0} \frac{1-\cos x}{x^2} = \lim_{x\to 0} \frac{2\sin^2 (x/2)}{x^2} = \frac{1}{2} \lim_{x\to 0} \left(\frac{\sin (x/2)}{x/2}\right)^2 = \frac{1}{2}.$$

On se sert souvent du symbole o (lire « o petit ») pour comparer les infiniment petits. Si en un point x_0 une fonction $\alpha(x)$ est un infiniment petit d'ordre supérieur à un infiniment petit $\beta(x)$, on note ceci :

$$\alpha(x) = o(\beta(x)).$$

Si des fonctions $\alpha(x)$ et $\beta(x)$ sont des infiniment petits en un point x_0, la fonction $\alpha(x)\beta(x)$ est d'un ordre de petitesse supérieur à celui de chaque facteur. En effet,

$$\lim_{x\to x_0} \frac{\alpha(x)\beta(x)}{\beta(x)} = \lim_{x\to x_0} \alpha(x) = 0,$$

donc $\alpha(x)\beta(x) = o(\beta(x))$, $\alpha(x)\beta(x) = o(\alpha(x))$.

Si $\alpha(x) \sim \alpha_1(x)$ et $\beta(x) \sim \beta_1(x)$ lorsque $x \to x_0$ et si $\lim\limits_{x\to x_0} \dfrac{\alpha(x)}{\beta(x)}$ existe, il en est de même de $\lim\limits_{x\to x_0} \dfrac{\alpha_1(x)}{\beta_1(x)}$ et de plus $\lim\limits_{x\to x_0} \dfrac{\alpha_1(x)}{\beta_1(x)} = \lim\limits_{x\to x_0} \dfrac{\alpha(x)}{\beta(x)}$. En effet,

$$\lim_{x\to x_0} \frac{\alpha_1(x)}{\beta_1(x)} = \lim_{x\to x_0} \left[\frac{\alpha_1(x)}{\alpha(x)} \frac{\alpha(x)}{\beta(x)} \frac{\beta(x)}{\beta_1(x)}\right] =$$

$$= \lim_{x\to x_0} \frac{\alpha_1(x)}{\alpha(x)} \lim_{x\to x_0} \frac{\alpha(x)}{\beta(x)} \lim_{x\to x_0} \frac{\beta(x)}{\beta_1(x)} =$$

$$= 1 \cdot \lim_{x\to x_0} \frac{\alpha(x)}{\beta(x)} \cdot 1 = \lim_{x\to x_0} \frac{\alpha(x)}{\beta(x)}.$$

Cette proposition simplifie le calcul des limites dans de nombreux cas.

Exemple. Calculer $\lim\limits_{x\to 0} \dfrac{\sin 5x}{x+x^3}$.

Solution. Puisque $\sin 5x \sim 5x$ et $x + x^3 \sim x$ lorsque $x \to 0$, on a

$$\lim_{x\to 0} \frac{\sin 5x}{x+x^3} = \lim_{x\to 0} \frac{5x}{x} = 5.$$

Les règles de comparaison des infiniment grands sont identiques.

Etudions quelques exemples.

1. Les fonctions $\alpha(x) = (1 + x)/x$ et $\beta(x) = 1/x$ sont des infiniment grands équivalents lorsque $x \to 0$, puisque

$$\lim_{x \to 0} \frac{\alpha(x)}{\beta(x)} = \lim_{x \to 0} (1 + x) = 1.$$

On dit encore que $\alpha(x)$ et $\beta(x)$ sont du même ordre de croissance lorsque $x \to 0$.

2. La fonction $\alpha(x) = x^2 + 4$ est un infiniment grand d'ordre inférieur à $\beta(x) = x^3 - 2$ lorsque $x \to \infty$ (admet un ordre de croissance inférieur), puisque

$$\lim_{x \to \infty} \frac{\alpha(x)}{\beta(x)} = \lim_{x \to \infty} \frac{x^2+4}{x^3-2} = \lim_{x \to \infty} \frac{1+4/x^2}{x-2/x^2} = \lim_{x \to \infty} \frac{1}{x} = 0.$$

3. Les infiniment grands $\alpha(x) = 2x^2 + 1$ et $\beta(x) = x^2 - 1$ sont du même ordre de croissance, puisque

$$\lim_{x \to \infty} \frac{2x^2+1}{x^2-1} = \lim_{x \to \infty} \frac{2+1/x^2}{1-1/x^2} = 2.$$

4. La fonction $\alpha(x) = x^4 + x + 1$ est un infiniment grand d'ordre deux par rapport à l'infiniment grand $\beta(x) = x^2 + 1$ lorsque $x \to \infty$, puisque

$$\lim_{x \to \infty} \frac{x^4+x+1}{(x^2+1)^2} = \lim_{x \to \infty} \frac{x^4+x+1}{x^4+2x^2+1} = \lim_{x \to \infty} \frac{1+1/x^3+1/x^4}{1+2/x^2+1/x^4} = 1.$$

§ 7. Notion de continuité d'une fonction

La notion de continuité d'une fonction est fondamentale en analyse mathématique.

1. Définition de la continuité d'une fonction. Soit $f(x)$ une fonction définie dans un voisinage d'un point x_0.

Définition 1. On dit que la fonction $f(x)$ est *continue au point* x_0 si elle est égale à sa limite en ce point, autrement dit si

$$\lim_{x \to x_0} f(x) = f(x_0) \text{ *}). \qquad (1)$$

Puisque $\lim_{x \to x_0} x = x_0$, la relation (1) peut être mise sous la forme

$$\lim_{x \to x_0} f(x) = f\left(\lim_{x \to x_0} x\right),$$

autrement dit, si une fonction est continue, les symboles f et lim sont permutables.

*) De la définition il s'ensuit que si une fonction est continue en x_0, elle est définie en ce point, c'est-à-dire que $f(x_0)$ existe. Remarquons que ceci n'était pas exigé dans la définition de la limite d'une fonction en x_0.

La continuité d'une fonction se définit de manière équivalente dans le langage des suites: on dit qu'une fonction $f(x)$ est continue en un point x_0 si pour toute suite $\{x_n\}$ de valeurs de x convergeant vers x_0, la suite $\{f(x_n)\}$ des valeurs respectives de la fonction $f(x)$ converge vers $f(x_0)$.

Formulons la définition de la continuité d'une fonction en termes de ε et δ.

Définition 2. On dit qu'une fonction $f(x)$ est *continue en un point* x_0 si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout x vérifiant la condition $|x - x_0| < \delta$ l'on ait $|f(x) - f(x_0)| < \varepsilon$.

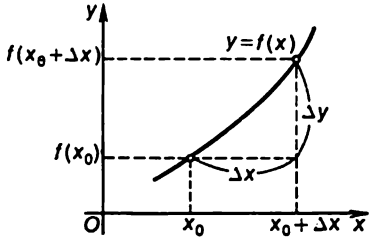


Fig. 50

L'équivalence de ces définitions est évidente.

En signes logiques la définition 2 s'écrit:

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in X,$$
$$|x - x_0| < \delta : |f(x) - f(x_0)| < \varepsilon.$$

Si $\lim\limits_{x \to x_0+} f(x) = f(x_0)$ (resp. $\lim\limits_{x \to x_0-} f(x) = f(x_0)$), la fonction $f(x)$ est dite *continue à droite* (resp. *à gauche*) au point x_0. Si la fonction $f(x)$ est continue à droite et à gauche en x_0, elle est *continue en* x_0. En effet, dans ce cas elle est égale à sa limite en ce point d'après le théorème 4.2.

Citons encore une définition de la continuité qui en fait est calquée sur la première.

Faisons passer $f(x_0)$ sous le signe lim dans (1). Les conditions $x \to x_0$ et $x - x_0 \to 0$ étant équivalentes, on a

$$\lim_{x - x_0 \to 0} [f(x) - f(x_0)] = 0. \qquad (2)$$

La différence $x - x_0$ s'appelle *accroissement de l'argument* x *en* x_0 et se note Δx, la différence $f(x) - f(x_0)$, *accroissement de la fonction en* x_0 correspondant à l'accroissement Δx de l'argument et se note Δy. Donc

$$\Delta x = x - x_0,\ \Delta y = f(x_0 + \Delta x) - f(x_0).$$

Signalons que si x_0 est fixe, l'accroissement Δy est une fonction de Δx. La signification géométrique des accroissements est évidente (fig. 50). Dans les nouvelles notations, l'égalité (2) devient

$$\lim_{\Delta x \to 0} \Delta y = 0. \qquad (3)$$

La relation (3) nous suggère encore une définition de la continuité d'une fonction.

Définition 3. On dit qu'une fonction $f(x)$ est *continue en un point* x_0 si son accroissement en ce point est un infiniment petit lorsque $\Delta x \to 0$.

La dernière définition est parfois plus commode à l'usage.

2. Opérations arithmétiques sur les fonctions continues.

Théorème 4.7. *Si* $f(x)$ *et* $g(x)$ *sont des fonctions continues en* x_0, *il en est de même des fonctions* $f(x) \pm g(x)$, $f(x)\,g(x)$ *et* $\frac{f(x)}{g(x)}$ (*dans le dernier cas sous réserve que* $g(x_0) \neq 0$).

Démonstration. Puisque les fonctions $f(x)$ et $g(x)$ continues en x_0 admettent en ce point des limites égales à $f(x_0)$ et $g(x_0)$, les fonctions $f(x) \pm g(x)$, $f(x)\,g(x)$ et $\frac{f(x)}{g(x)}$ admettent aussi d'après le théorème 4.3 des limites qui sont respectivement égales à $f(x_0) \pm g(x_0)$, $f(x_0)\,g(x_0)$ et $\frac{f(x_0)}{g(x_0)}$. Mais ces quantités sont égales aux valeurs prises par ces fonctions en x_0. Donc les fonctions $f(x) \pm g(x)$, $f(x)\,g(x)$ et $\frac{f(x)}{g(x)}$ sont continues en x_0 aux termes de la définition 1. ■

§ 8. Continuité de certaines fonctions élémentaires

Une importante propriété des fonctions élémentaires est leur continuité en chaque point au voisinage duquel elles sont définies. Vérifions ce fait sur l'exemple de quelques fonctions en utilisant la définition de la continuité d'une fonction en un point et le théorème 4.7.

1. Continuité des fonctions rationnelles. Un exemple élémentaire de fonction continue en tout point x_0 de la droite numérique est la fonction constante $f(x) = C$. En effet, $\lim\limits_{x \to x_0} f(x) = C = f(x_0)$ (cf. exemple 1, § 2), autrement dit la fonction constante est continue en tout point de la droite numérique.

Est continue aussi en tout point de la droite numérique la fonction $f(x) = x$, puisque $\lim\limits_{x \to x_0} x = x_0 = f(x_0)$ (cf. exemple 2, § 2), c'est-à-dire que la limite de cette fonction est égale à sa valeur en ce point. De ce qui précède et du théorème 4.7 il résulte que les fonctions $x^2 = x \cdot x$, $x^3 = x^2 \cdot x$, $x^4 = x^3 \cdot x$, ... , $x^n = x^{n-1} \cdot x$ (n est un entier naturel) sont continues en tout point x_0.

On sait que la fonction $f(x) = x^n$ s'appelle fonction puissance et la fonction

$$P(x) = C_0 x^n + C_1 x^{n-1} + C_2 x^{n-2} + \ldots + C_{n-1} x + C_n,$$

où $n \geqslant 0$ est un entier et $C_0, C_1, \ldots, C_n$ des nombres arbitraires, fonction polynôme.

Chaque terme

$$C_0x^n,\ C_1x^{n-1},\ C_2x^{n-2},\ \ldots,\ C_n$$

est le produit de deux fonctions continues (une fonction constante et une fonction puissance). Il est continu en tout point x d'après le théorème 4.7. Le polynôme $P(x)$ est donc la somme de fonctions continues en tout point x et par suite est continu en tout point x.

La fonction rationnelle

$$R(x) = \frac{P(x)}{Q(x)}$$

où $P(x)$ et $Q(x)$ sont des fonctions polynômes, est continue en tout point en lequel $Q(x) \neq 0$ (partout sauf aux zéros de $Q(x)$) en tant que quotient de fonctions continues.

Ainsi la fonction $R(x) = \frac{3x^2+7x-1}{x^2-1}$ est continue en tous les points x distincts de ± 1.

2. Continuité des fonctions circulaires. Considérons les fonctions circulaires $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{cotg} x$, $\sec x$ et $\operatorname{cosec} x$. Montrons que la fonction $\sin x$ est continue en tout point x. Utilisons pour cela la définition 3 de la continuité d'une fonction. En faisant subir l'accroissement Δx à x, on obtient l'accroissement

$$\Delta y = \sin(x + \Delta x) - \sin x$$

ou

$$\Delta y = 2\cos\left(x + \frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2}.$$

En faisant tendre Δx vers 0, on trouve

$$\lim_{\Delta x \to 0} \Delta y = 2 \lim_{\Delta x \to 0} \left[\cos\left(x + \frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2}\right] = 0,$$

puisque $|\cos(x + \Delta x/2)| \leqslant 1$,

$$\lim_{\Delta x \to 0} \sin\frac{\Delta x}{2} = \lim_{\Delta x \to 0} \frac{\sin(\Delta x/2)}{\Delta x/2} \lim_{\Delta x \to 0} \frac{\Delta x}{2} = \frac{1}{2} \cdot 1 \cdot \lim_{\Delta x \to 0} \Delta x = 0$$

et le produit d'une fonction bornée par un infiniment petit est un infiniment petit. La fonction $\sin x$ est donc continue en tout point x.

La continuité de la fonction $\cos x$ en tout point x se prouve de façon analogue.

La continuité des fonctions $\sin x$ et $\cos x$ entraîne en vertu du théorème 4.7 celle des fonctions $\operatorname{tg} x = \frac{\sin x}{\cos x}$, $\sec x = \frac{1}{\cos x}$ en tous les points où $\cos x \neq 0$, c'est-à-dire en tous les points excepté les points $x = \frac{\pi}{2} + n\pi$, et celle des fonctions $\operatorname{cotg} x =$

$= \frac{\cos x}{\sin x}$ et cosec $x = \frac{1}{\sin x}$ en tous les points excepté les points $x = n\pi$ $(n = 0, \pm 1, \pm 2, \ldots)$.

3. Continuité de la fonction $f(x) = |x|$. La fonction $f(x) = |x|$ dont la courbe est représentée sur la figure 51 est définie et continue en tout point de la droite numérique. En effet, elle est continue en tout point de la demi-droite $]0, +\infty[$, puisque $f(x) = x$ pour $x > 0$ (cf. n° 1). Elle est aussi continue en tout point de la demi-droite $]-\infty, 0[$, puisque $f(x) = -x$ pour $x < 0$ et on peut la représenter comme le produit de deux fonctions continues, les fonctions -1 et x, et lui appliquer le théorème 4.7 de continuité d'un produit. Pour établir la continuité de la fonction $|x|$ en $x = 0$, calculons les limites à droite et à gauche en $x = 0$:

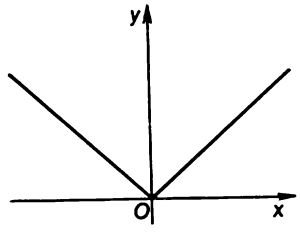


Fig. 51

$$\lim_{x \to 0-} |x| = \lim_{x \to 0-} (-x) = -\lim_{x \to 0-} x = 0; \quad \lim_{x \to 0+} |x| = \lim_{x \to 0+} x = 0.$$

La fonction $|x|$ admet en $x = 0$ des limites à droite et à gauche qui sont égales entre elles et à sa valeur en $x = 0$. Il en résulte que la fonction $|x|$ est continue en $x = 0$ et par suite en tous les points de la droite numérique.

Donc les fonctions envisagées sont continues en tout point au voisinage duquel elles sont définies. Le théorème 4.7 de continuité de la somme, de la différence, du produit et du quotient nous permet d'affirmer que les fonctions obtenues à l'aide d'un nombre fini d'opérations arithmétiques sur des fonctions continues sont continues en chaque point au voisinage duquel elles sont définies.

On dira qu'une fonction $f(x)$ est continue sur un intervalle $]a, b[$ si elle l'est en tout point de cet intervalle ; continue sur un segment $[a, b]$ si elle l'est sur l'intervalle $]a, b[$, en a à droite et en b à gauche, c'est-à-dire que

$$\lim_{x \to a+} f(x) = f(a) \quad \text{et} \quad \lim_{x \to b-} f(x) = f(b).$$

§ 9. Classification des points de discontinuité d'une fonction

1. Définition et classification des points de discontinuité d'une fonction.

Définition. On dit qu'un point x_0 est un *point de discontinuité d'une fonction* $f(x)$ si $f(x)$ n'y est pas continue.

Les discontinuités des fonctions sont de deux espèces.

Discontinuité de première espèce. On dit qu'une fonction $f(x)$ présente une *discontinuité de première espèce en un point* x_0 ou encore que x_0 est un *point de discontinuité de première espèce* si $f(x)$ admet en ce point des limites à droite et à gauche qui ne sont pas égales:

$$\lim_{x \to x_0+} f(x) \neq \lim_{x \to x_0-} f(x).$$

Exemple. La fonction $f(x) = \operatorname{sgn} x$ présente une discontinuité de première espèce (cf. fig. 46) en $x = 0$, puisque $\lim\limits_{x \to 0+} \operatorname{sgn} x = 1$, $\lim\limits_{x \to 0-} \operatorname{sgn} x = -1$.

Discontinuité de deuxième espèce. On dit qu'une fonction $f(x)$ présente une *discontinuité de deuxième espèce en un point* x_0 ou encore que x_0 est un *point de discontinuité de deuxième espèce* si l'une au moins de ses dérivées à droite et à gauche en x_0 n'existe pas ou est infinie.

Exemple. Le point $x = 0$ est un point de discontinuité de deuxième espèce pour la fonction $f(x) = \frac{1}{x}$ (cf. fig. 48), puisque $\lim\limits_{x \to 0+} \frac{1}{x} = +\infty$ et $\lim\limits_{x \to 0-} \frac{1}{x} = -\infty$.

2. Fonctions continues par morceaux. On dit qu'une fonction $f(x)$ est *continue par morceaux sur un intervalle* $[a, b]$ si elle est continue en tous les points intérieurs de $[a, b]$ à l'exception éventuellement d'un nombre fini de points en lesquels elle présente des discontinuités de première espèce, et si elle admet une limite à droite en a et une à gauche en b.

Une fonction est *continue par morceaux sur la droite numérique* tout entière si elle l'est sur tout segment de cette droite.

Exemple. La fonction $f(x) = [x]$ est continue par morceaux sur toute la droite numérique. Le graphique de la fonction $f(x) = [x]$ est représenté sur la figure 52. La fonction $[x]$ est continue à droite et discontinue à gauche aux points $x = n$ ($n = 0, \pm 1, \pm 2, \ldots$). En tous les autres points elle est continue aussi bien à droite qu'à gauche.

§ 10. Propriétés fondamentales des fonctions continues

1. Théorème de stabilité du signe d'une fonction continue.

Théorème 4.8. *Si une fonction $f(x)$ est continue en un point x_0 en lequel elle ne s'annule pas, il existe un δ-voisinage de x_0 dans lequel elle est du même signe que $f(x_0)$.*

Démonstration. Supposons que $f(x_0) > 0$ (fig. 53). Alors en vertu de la définition 2 de la continuité d'une fonction, pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que $|f(x) - f(x_0)| < \varepsilon$

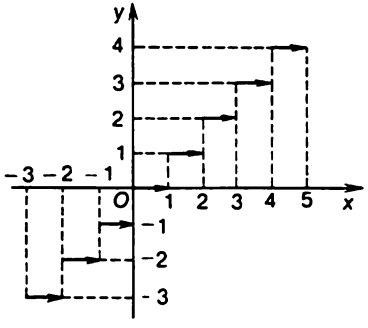


Fig. 52

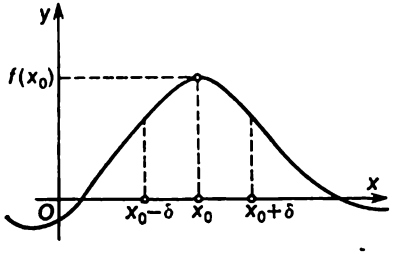


Fig. 53

pour tout x vérifiant la condition $|x - x_0| < \delta$ ou, ce qui est équivalent, tel que

$$f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon \qquad (1)$$

pour tout $x \in]x_0 - \delta, x_0 + \delta[$. Prenons $\varepsilon = f(x_0)$. De l'inégalité de gauche (1) on déduit que $f(x) > 0$ pour tout $x \in]x_0 - \delta, x_0 + \delta[$. C.Q.F.D.

Si $f(x_0) < 0$, considérons la fonction $-f(x)$. Puisque $-f(x_0) > 0$, d'après ce qui vient d'être prouvé, il existe un δ-voisinage du point x_0 dans lequel $-f(x) > 0$ et par suite $f(x) < 0$. ■

2. Passage d'une fonction continue par toute valeur intermédiaire. Considérons le théorème de changement de signe d'une fonction continue.

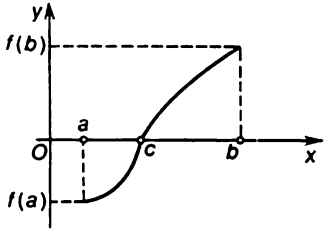


Fig. 54

Théorème 4.9 (premier théorème de Bolzano-Cauchy) *). *Si une fonction $f(x)$ est continue sur un intervalle $[a, b]$ aux bornes duquel elle prend des valeurs de signes contraires, il existe un point $c \in]a, b[$ en lequel $f(c) = 0$.*

Démonstration. Supposons pour fixer les idées que $f(a) < 0$ et $f(b) > 0$ (fig. 54). Divisons le segment $[a, b]$ en deux.

*) Bolzano Bernard (1781-1848), mathématicien tchèque.

Si la fonction $f(x)$ s'annule au milieu du segment $[a, b]$, le théorème est prouvé. Sinon choisissons celui des segments obtenus aux bornes duquel la fonction $f(x)$ prend des valeurs de signes contraires, désignons-le par $[a_1, b_1]$ et partageons-le en deux. Si la fonction $f(x)$ s'annule au milieu du segment $[a_1, b_1]$, le théorème est prouvé. Sinon choisissons celui des segments obtenus aux bornes duquel $f(x)$ prend des valeurs de signes contraires et désignons-le par $[a_2, b_2]$. Si l'on poursuit cette procédure, soit la fonction $f(x)$ s'annulera au milieu d'un segment $[a_k, b_k]$ et le théorème sera prouvé, soit on obtiendra une suite

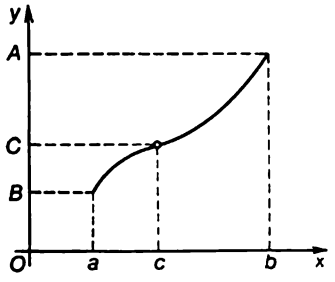


Fig. 55

$$[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \ldots$$
$$\ldots \supset [a_n, b_n] \supset \ldots$$

de segments emboîtés tels que $b_n - a_n = (b - a)/2^n \to 0$ lorsque $n \to \infty$ et aux bornes de chacun desquels $f(x)$ prend des valeurs de signes contraires.

D'après le théorème 2.13 des segments emboîtés, il existe un point c intérieur à tous ces segments. Montrons que $f(c) = 0$. En effet, si l'on admet que $f(c) > 0$, il existe d'après le théorème 4.8 un voisinage du point c dans lequel $f(x) > 0$. Pour n assez grand, le segment $[a_n, b_n]$ sera contenu dans ce voisinage et l'on aura $f(x) > 0$ sur $[a_n, b_n]$. Ce qui contredit le fait que $f(x)$ prend des valeurs de signes contraires aux bornes de $[a_n, b_n]$. On démontre de même que $f(c)$ ne peut être strictement inférieur à 0. Donc $f(c) = 0$. De plus il est évident que $c \in]a, b[$. ■

Ce théorème admet une interprétation géométrique simple: toute courbe continue coupe l'axe des abscisses en passant d'un demi-plan (supérieur ou inférieur) à l'autre.

Considérons le théorème de passage d'une fonction continue par toute valeur intermédiaire.

Théorème 4.10 (deuxième théorème de Bolzano-Cauchy). *Si une fonction $f(x)$ est continue sur un intervalle $[a, b]$, $f(a) = A$, $f(b) = B$ et C est une valeur comprise entre A et B, alors il existe un point $c \in [a, b]$ tel que $f(c) = C$.*

En d'autres termes, en passant d'une valeur à une autre une fonction continue prend toutes les valeurs intermédiaires.

Démonstration. Supposons pour fixer les idées que $A < B$ et $A < C < B$ (fig. 55 où il faut inverser la position de A et B). Considérons la fonction auxiliaire

$$\varphi(x) = f(x) - C.$$

Cette fonction est continue sur $[a, b]$ et prend des valeurs de signes contraires aux bornes:

$$\varphi(a) = f(a) - C = A - C < 0,$$
$$\varphi(b) = f(b) - C = B - C > 0.$$

D'après le théorème 4.9, il existe un point $c \in]a, b[$ tel que $\varphi(c) = f(c) - C = 0$. D'où $f(c) = C$. ■

Corollaire. *Si une fonction $f(x)$ est définie et continue sur un intervalle X, son ensemble des valeurs Y est aussi un intervalle.*

Avant de prouver ce corollaire, introduisons la notion de bornes d'une fonction. Soient $y = f(x)$ une fonction définie sur X; Y, son ensemble des valeurs. Si l'ensemble Y est majoré (resp. minoré), il admet une borne supérieure (resp. inférieure). La borne supérieure (resp. inférieure) de l'ensemble Y s'appelle *borne supérieure* (resp. *inférieure*) *de la fonction $y = f(x)$ sur l'ensemble X* et se note $\sup\limits_X f(x)$ (resp. $\inf\limits_X f(x)$).

La définition de la borne supérieure (resp. inférieure) peut encore être formulée dans les termes suivants:

On dit qu'un nombre M (resp. m) est la borne supérieure (resp. inférieure) d'une fonction $y = f(x)$ sur un ensemble X si sont remplies les deux conditions suivantes:

1) $f(x) \leqslant M$ (resp. $f(x) \geqslant m$) pour tout $x \in X$;

2) pour tout nombre $M' < M$ (resp. $m' > m$) il existe un point $x' \in X$ tel que $f(x') > M'$ (resp. $f(x') < m'$).

La première de ces conditions exprime que M (resp. m) est un majorant (resp. minorant) de la fonction $y = f(x)$ sur X, la deuxième que M (resp. m) est le plus petit (resp. le plus grand) des majorants (resp. minorants) de la fonction, c'est-à-dire la borne supérieure (resp. inférieure).

Si l'ensemble Y n'est pas majoré (resp. minoré), on écrit $\sup\limits_X f(x) = +\infty$ (resp. $\inf\limits_X f(x) = -\infty$). Dans ce cas pour tout nombre A il existe un point $x' \in X$ tel que $f(x') > A$ (resp. $f(x') < A$).

Prouvons maintenant le corollaire du théorème 4.10.

Démonstration. Soient $m = \inf\limits_X f(x)$, $M = \sup\limits_X f(x)$. Prenons trois valeurs y_1, y et y_2 de Y telles que

$$m \leqslant y_1 < y < y_2 \leqslant M$$

(si $M = +\infty$ (resp. $m = -\infty$), on prend $y_2 < M$ (resp. $m < y_1$)). L'existence de telles valeurs de la fonction $f(x)$ résulte de la définition des bornes d'une fonction. D'après le théorème 4.10 des valeurs intermédiaires d'une fonction continue, il existe un point x tel que $f(x) = y$. Donc, l'ensemble Y est un intervalle (fini ou infini) d'extrémités m et M, ouvert ou fermé selon les cas. ■

3. Théorème de limitation d'une fonction continue sur un segment. On rappelle qu'une fonction $f(x)$ est bornée sur un intervalle $[a, b]$ s'il existe un nombre $M > 0$ tel que pour tout $x \in [a, b]$ on a $|f(x)| \leqslant M$ ou $-M \leqslant f(x) \leqslant M$, c'est-à-dire que le graphique de la fonction $f(x)$ est situé dans la région intérieure aux droites $y = M$ et $y = -M$ (fig. 56).

Théorème 4.11 (premier théorème de Weierstrass) *). *Une fonction $f(x)$ définie et continue sur un intervalle $[a, b]$ est bornée sur cet intervalle.*

Prouvons préalablement le lemme suivant.

Lemme. *Une fonction $f(x)$ continue en un point x_0 est bornée dans un voisinage de ce point.*

Démonstration. Soit $\varepsilon = 1$; d'après la définition 2 de la continuité d'une fonction en un point, pour cette valeur de ε il existe un $\delta > 0$ tel que $|f(x) - f(x_0)| < 1$ dans un δ-voisinage de x_0. Cette inégalité nous donne $|f(x)| = |(f(x) - f(x_0)) + f(x_0)| \leqslant |f(x) - f(x_0)| + |f(x_0)| < 1 + |f(x_0)|$, c'est-à-dire que $|f(x)| < M$, où $M = 1 + |f(x_0)|$. D'où il résulte que la fonction $f(x)$ est bornée dans un δ-voisinage du point x_0. ■

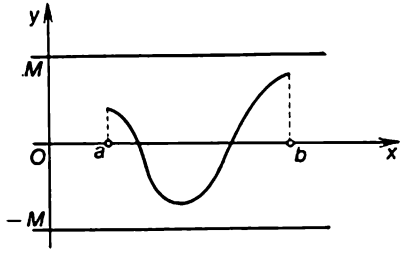


Fig. 56

Démonstration du théorème. Supposons par absurde que la fonction $f(x)$ n'est pas bornée sur l'intervalle $[a, b]$. Divisons l'intervalle $[a, b]$ en deux. Sur l'un au moins de ces segments que l'on désignera par $[a_1, b_1]$ la fonction $f(x)$ n'est pas bornée (sinon elle serait bornée sur $[a, b]$). Divisons $[a_1, b_1]$ en deux et désignons par $[a_2, b_2]$ le segment sur lequel $f(x)$ n'est pas bornée. En poursuivant cette procédure, on obtient une suite

$$[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \ldots \supset [a_n, b_n] \supset \ldots$$

de segments emboîtés sur chacun desquels $f(x)$ n'est pas bornée et tels que $b_n - a_n = \frac{b-a}{2^n} \to 0$ lorsque $n \to \infty$.

D'après le théorème 2.13 des segments emboîtés, il existe un point c intérieur à tous ces segments. La fonction $f(x)$ est par hypothèse définie et continue en c, donc en vertu du lemme prouvé elle est bornée dans un voisinage du point c. A partir d'un certain n assez grand ce voisinage contiendra des segments $[a_n, b_n]$ sur lesquels la fonction $f(x)$ est aussi bornée. Ceci contredit le

*) Weierstrass Karl (1815-1897), mathématicien allemand.

fait que $f(x)$ n'est pas bornée sur chaque segment emboîté. Cette contradiction prouve le théorème. ■

Remarque. Le théorème est mis en défaut si l'on remplace l'intervalle $[a, b]$ par l'intervalle $]a, b[$. Par exemple, la fonction $f(x) = \frac{1}{x}$ est continue sur $]0, 1[$ mais pas bornée, puisque $\lim\limits_{x \to 0+} \frac{1}{x} = +\infty$. La démonstration du théorème ne « passe pas » là où on affirme que la fonction est définie et continue au

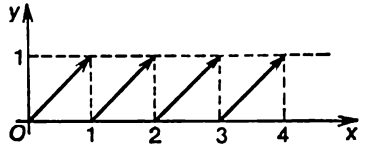


Fig. 57

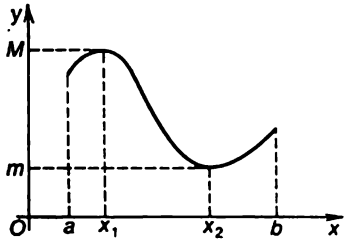


Fig. 58

point c. Si l'intervalle est ouvert, le point c risque d'être confondu avec l'une de ses extrémités et la fonction $f(x)$ ne sera pas définie et continue en c.

4. Théorème d'atteinte des bornes d'une fonction continue sur un intervalle. Lorsque les bornes d'une fonction sont des valeurs de cette fonction, on dit qu'elle atteint ses bornes. Tout ensemble ne contient pas ses bornes (cf. formule (1), chap. I, théorème 1.1). L'exemple suivant montre qu'une fonction n'atteint pas toujours ses bornes.

Soit donnée sur l'intervalle $[0, b]$, $b \geqslant 1$, la fonction $f(x) = x - [x]$ dont le graphique est représenté sur la figure 57. Son ensemble de valeurs est l'intervalle $[0, 1[$. Cette fonction possède une borne supérieure égale à 1 et une borne inférieure égale à 0. Il est évident que $f(x)$ prend la valeur 0 mais pas la valeur 1. Donc elle atteint sa borne inférieure mais pas sa borne supérieure.

Voyons sous quelle condition une fonction atteint ses bornes.

Théorème 4.12 (deuxième théorème de Weierstrass). *Une fonction $f(x)$ continue sur un intervalle $[a, b]$ atteint ses bornes sur cet intervalle, autrement dit il existe des points x_1 et x_2 de $[a, b]$ en lesquels* (fig. 58)

$$f(x_1) = M = \sup_{[a, b]} f(x), \quad f(x_2) = m = \inf_{[a, b]} f(x).$$

Démonstration. Sur l'intervalle $[a, b]$ la fonction $f(x)$ est continue, donc bornée d'après le théorème 4.11. Par conséquent

elle possède une borne supérieure M et une borne inférieure m d'après le théorème 1.1.

Montrons que la fonction $f(x)$ atteint la borne M, c'est-à-dire qu'il existe un point $x_1 \in [a, b]$ tel que $f(x_1) = M$. Raisonnons par l'absurde. Supposons que la fonction $f(x)$ ne prend la valeur M en aucun point de $[a, b]$. Pour tout $x \in [a, b]$ on aura alors $f(x) < M$. Considérons la fonction auxiliaire

$$F(x) = \frac{1}{M - f(x)}$$

strictement positive sur $[a, b]$.

D'après le théorème 4.7, la fonction $F(x)$ est continue en tant que quotient de deux fonctions continues. Elle est alors bornée d'après le théorème 4.11, ce qui veut dire qu'il existe un nombre $\mu > 0$ tel que pour tout $x \in [a, b]$

$$F(x) = \frac{1}{M - f(x)} \leqslant \mu,$$

d'où

$$f(x) \leqslant M - \frac{1}{\mu}.$$

Donc le nombre $M - \frac{1}{\mu}$ qui est strictement inférieur à M est la borne supérieure de $f(x)$ sur $[a, b]$. Or ceci contredit le fait que M est la borne supérieure. Cette contradiction prouve qu'il existe un point $x_1 \in [a, b]$ en lequel $f(x_1) = M$.

On démontre de façon analogue que la fonction $f(x)$ atteint sa borne inférieure m sur $[a, b]$. ■

Remarque 1. Maintenant qu'on a démontré qu'une fonction $f(x)$ continue sur un intervalle $[a, b]$ atteint ses bornes supérieure M et inférieure m sur cet intervalle, on peut appeler M maximum et m minimum de $f(x)$ sur $[a, b]$ et énoncer le théorème 4.12 sous la forme suivante: *une fonction continue sur un intervalle prend son maximum et son minimum sur cet intervalle.*

Remarque 2. La différence entre le maximum et le minimum d'une fonction $f(x)$ continue sur un intervalle $[a, b]$ s'appelle *oscillation* de $f(x)$ sur $[a, b]$ et se note $\omega : \omega = M - m$, où

$$M = \max_{[a, b]} f(x), \quad m = \min_{[a, b]} f(x).$$

5. Notion de continuité uniforme d'une fonction. La continuité uniforme est une très importante propriété des fonctions continues sur un intervalle. Elle est largement utilisée dans la démonstration de nombreux théorèmes fondamentaux.

Soit $f(x)$ une fonction continue sur un intervalle X et soit $x_0 \in X$. La fonction $f(x)$ étant continue au point x_0, en vertu de la définition 2 de la continuité, pour tout $\varepsilon > 0$ il existe un $\delta > 0$

tel que $|f(x) - f(x_0)| < \varepsilon$ pour $|x - x_0| < \delta$. Il est clair que δ dépend de ε et aussi de x_0. Lorsque x_0 varie dans les limites de l'intervalle envisagé (à ε constant), le nombre δ sera différent d'un x_0 à l'autre. Plus le graphique est « abrupt » au voisinage de x_0, plus le nombre δ correspondant à ce point est petit (fig. 59).

Donc pour ε donné, à tout point x de cet intervalle correspond un $\delta > 0$. Si les points x étaient en nombre fini, parmi les δ correspondants on pourrait choisir le plus petit et ce δ dépendrait

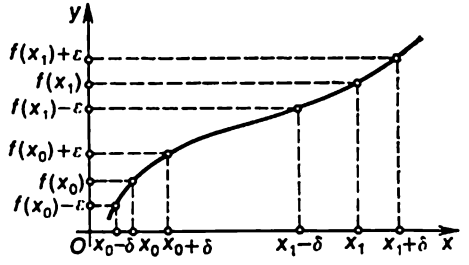


Fig. 59

seulement de ε et « conviendrait » pour tous les x. Si les nombres x sont en quantité infinie, cette procédure ne passe pas en général, car à ces points correspond une infinité de nombres δ parmi lesquels figurent des nombres aussi petits que l'on veut.

Il se pose la question de savoir s'il existe des fonctions continues, définies sur des intervalles tels que pour tout $\varepsilon > 0$ on trouverait un $\delta > 0$ indépendant de x, c'est-à-dire que δ serait commun à tous les x de l'intervalle envisagé. Ceci nous conduit à la notion de continuité uniforme d'une fonction.

Définition. On dit qu'une fonction $f(x)$ est *uniformément continue sur un intervalle* X si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout couple de points x', $x'' \in X$ vérifiant l'inégalité $|x'' - x'| < \delta$ l'on ait $|f(x'') - f(x')| < \varepsilon$.

En signes logiques cette définition s'écrit:

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x',\ x'' \in X;\ |x'' - x'| < \delta:\ |f(x'') - f(x')| < \varepsilon.$$

Par définition, δ dépend seulement de ε et est commun à tous les x', x'' de l'intervalle X. Il résulte de la définition qu'une fonction uniformément continue sur X est continue sur X.

Le théorème suivant établit la réciproque de cette proposition.

6. Théorème de continuité uniforme d'une fonction.

Théorème 4.13 (Cantor) *). *Une fonction $f(x)$ continue sur un intervalle $[a, b]$ est uniformément continue sur cet intervalle.*

*) Cantor Georg (1845-1918), mathématicien allemand, fondateur de la théorie des ensembles.

Démonstration. Montrons tout d'abord que si $f(x)$ est continue sur $[a, b]$, alors pour tout $\varepsilon > 0$ le segment $[a, b]$ peut être subdivisé en un nombre fini de segments soit deux à deux disjoints, soit ayant deux à deux une borne commune et tels que sur chacun d'eux $|f(x'') - f(x')| < \varepsilon$ quels que soient x' et x''.

Supposons par absurde qu'il existe un $\varepsilon > 0$ pour lequel est impossible une telle subdivision du segment $[a, b]$. Divisons $[a, b]$ en deux et désignons par $[a_1, b_1]$ celui des deux segments pour lequel une telle subdivision est impossible. Divisons le segment $[a_1, b_1]$ en deux et désignons par $[a_2, b_2]$ celui des deux segments pour lequel une telle subdivision est impossible. En poursuivant cette procédure, on obtient une suite de segments emboîtés

$$[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \ldots \supset [a_n, b_n] \supset \ldots$$

tels qu'aucun d'entre eux ne peut être subdivisé en un nombre fini de segments sur chacun desquels $|f(x'') - f(x')| < \varepsilon$ quels que soient x' et x''. D'après le théorème 2.13 des segments emboîtés, il existe un point c intérieur à tous ces segments. La fonction $f(x)$ étant continue en c, pour la valeur de ε envisagée il existe un δ tel que $|f(x) - f(c)| < \frac{\varepsilon}{2}$ pour tout x pris dans le δ-voisinage de c. Pour deux points quelconques x' et x'' du δ-voisinage de c on aura

$$|f(x'') - f(x')| = |(f(x'') - f(c)) + (f(c) - f(x'))| \leqslant$$
$$\leqslant |f(x'') - f(c)| + |f(c) - f(x')| < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

c'est-à-dire que

$$|f(x'') - f(x')| < \varepsilon.$$

Le δ-voisinage du point c contiendra le segment $[a_n, b_n]$ pour n assez grand et, par suite, pour deux points quelconques x' et x'' de ce segment on aura $|f(x'') - f(x')| < \varepsilon$, ce qui contredit le choix de la suite des segments emboîtés.

Passons maintenant à la démonstration du théorème. D'après ce qui vient d'être prouvé, pour tout $\varepsilon > 0$ il existe une subdivision de $[a, b]$ en un nombre fini de segments sur chacun desquels l'écart entre deux valeurs quelconques de la fonction $f(x)$ est strictement inférieur à $\frac{\varepsilon}{2}$ en valeur absolue. Désignons par δ la longueur du plus petit segment de la subdivision et considérons deux points quelconques, x' et x'', du segment $[a, b]$ tels que $|x'' - x'| < \delta$. Deux cas sont possibles : 1) x' et x'' appartiennent à un même segment ; 2) x' et x'' appartiennent à deux segments voisins. Dans le premier cas, $|f(x'') - f(x')| < \frac{\varepsilon}{2}$; dans le deuxième en désignant par x_0 la borne commune de ces deux

segments, on obtient

$$|f(x'') - f(x')| = |(f(x'') - f(x_0)) + (f(x_0) - f(x'))| \leqslant$$
$$\leqslant |f(x'') - f(x_0)| + |f(x_0) - f(x')| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Donc, pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout couple de points x' et x'' du segment $[a, b]$ tels que $|x'' - x'| < \delta$ on a $|f(x'') - f(x')| < \varepsilon$. ■

Corollaire. *Si $f(x)$ est une fonction continue sur un intervalle $[a, b]$, pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que si l'on subdivise*

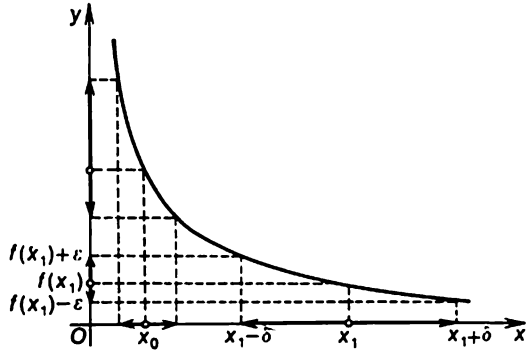


Fig. 60

$[a, b]$ arbitrairement en un nombre fini de segments de longueur strictement inférieure à δ, l'oscillation ω de $f(x)$ sur chacun d'eux sera strictement inférieure à ε.

Démonstration. En effet, la fonction $f(x)$ est uniformément continue sur $[a, b]$ d'après le théorème prouvé. Donc pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout couple x', x'' de $[a, b]$ vérifiant l'inégalité $|x'' - x'| < \delta$ on a $|f(x'') - f(x')| < \varepsilon$. Subdivisons arbitrairement le segment $[a, b]$ en un nombre fini de segments de longueur strictement inférieure à δ. La fonction $f(x)$ étant continue sur $[a, b]$, sur chaque sous-segment de la subdivision on peut exhiber des points x' et x'' tels que $f(x') = m$ et $f(x'') = M$, où m et M sont les bornes inférieure et supérieure de $f(x)$ sur le sous-segment envisagé. Comme $|x'' - x'| < \delta$, on a $|f(x'') - f(x')| < \varepsilon$. Or $f(x'') - f(x') = M - m = \omega$, donc $\omega < \varepsilon$. ■

Remarque. Le théorème est mis en défaut si l'on remplace le segment $[a, b]$ par un intervalle ouvert ou semi-ouvert.

Exemple. Considérons la fonction $f(x) = \frac{1}{x}$ sur l'intervalle $]0, 1[$. Cette fonction est continue mais pas uniformément continue sur $]0, 1[$. Ceci résulte du fait que pour tout $\varepsilon > 0$ fixe, quelque

petit que soit $\delta > 0$, on pourra toujours exhiber deux points x' et x'' assez proches de 0 tels que $|x'' - x'| < \delta$ et $|f(x'') - f(x')| > \varepsilon$ (fig. 60).

§ 11. Notion de fonction composée

Définition. Si $z = \varphi(x)$ est une fonction d'ensemble de définition X et d'ensemble des valeurs Z et $y = f(z)$ une fonction d'ensemble de définition Z, la fonction $y = f[\varphi(x)]$ s'appelle *fonction composée* *) de x et la variable z, *variable intermédiaire* de la fonction composée.

Exemple. La fonction $y = \sin x^2$ est une fonction composée définie sur la droite numérique tout entière, puisque $y = f(z) = \sin z$, $z = \varphi(x) = x^2$.

Théorème 4.14. *Si $z = \varphi(x)$ est une fonction continue en x_0 et $y = f(z)$, une fonction continue au point $z_0 = \varphi(x_0)$, la fonction composée $y = f[\varphi(x)]$ est continue en x_0.*

Démonstration. Soit dans X une suite de points $\{x_n\}$ convergeant vers x_0. La continuité de la fonction $z = \varphi(x)$ en x_0 nous permet d'écrire: $\lim\limits_{n\to\infty} z_n = \lim\limits_{n\to\infty} \varphi(x_n) = \varphi(x_0) = z_0$, c'est-à-dire que la suite $\{\varphi(x_n)\}$ converge vers z_0. La fonction $f(z)$ étant continue en z_0, on a $\lim\limits_{n\to\infty} f(z_n) = f(z_0)$, c'est-à-dire que $\lim\limits_{n\to\infty} f[\varphi(x_n)] = f[\varphi(x_0)]$. Donc la limite de la fonction $f[\varphi(x)]$ en x_0 est égale à sa valeur en ce point. Ce qui exprime que la fonction composée $f[\varphi(x)]$ est continue en x_0. ■

Exemple. Montrer que la fonction $y = \sin x^2$ est continue en $x = 0$.

Solution. La fonction $z = x^2$ est continue en $x = 0$, la fonction $y = \sin z$, en $z = 0$, donc d'après le théorème prouvé la fonction composée $y = \sin x^2$ est continue en $x = 0$.

§ 12. Notion de fonction réciproque

1. Définition de la fonction réciproque. On dira qu'une fonction $f(x)$ est *croissante* (resp. *décroissante*) sur un ensemble X si pour tout couple de points $x_1, x_2 \in X$ tels que $x_1 < x_2$ on a $f(x_1) \leqslant f(x_2)$ (resp. $f(x_1) \geqslant f(x_2)$).

Les fonctions croissantes et décroissantes sont dites *monotones*.

On dit qu'une fonction $f(x)$ est *strictement croissante* (resp. *strictement décroissante*) sur un ensemble si pour tout couple de

*) On se sert aussi du terme équivalent « fonction de fonction ».

points x_1, $x_2 \in X$ tels que $x_1 < x_2$ on a $f(x_1) < f(x_2)$ (resp. $f(x_1) > f(x_2)$).

Les fonctions strictement croissantes et strictement décroissantes sont dites *strictement monotones.*

Exemples. 1. La fonction $f(x) = \operatorname{sgn} x$ est croissante sur toute la droite numérique.

2. La fonction $f(x) = x$ est strictement croissante sur toute la droite numérique.

Introduisons maintenant la notion de fonction réciproque.

Définition. Soient X et Y des ensembles et soit donnée une fonction f, c'est-à-dire un ensemble de couples de nombres $(x; y)$ $(x \in X, y \in Y)$ tels que chaque nombre x fait partie d'un couple et d'un seul et chaque nombre y, d'au moins un couple. Si dans chaque couple de cet ensemble on permute x et y, on obtient un ensemble de couples $(y; x)$ qui s'appelle *fonction réciproque de f.*

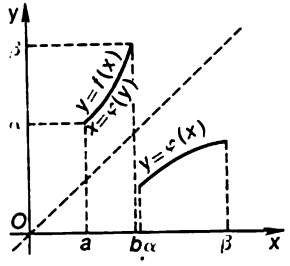


Fig. 61

La fonction réciproque sera désignée par le symbole $x = \varphi(y)$.

Signalons que la fonction réciproque n'est pas généralement une fonction, car chaque nombre y peut figurer dans un comme dans plusieurs couples. Par exemple, la fonction $x = y$ réciproque de la fonction $y = x$ est monovalente (chaque nombre y est élément d'un seul couple) ; la fonction $x = \pm \sqrt{y}$, réciproque de la fonction $y = x^2$, est bivalente (chaque nombre y est élément de deux couples) ; la fonction $y = \sin x$, réciproque de la fonction $x = \arcsin y$, est multivalente (chaque nombre y est élément d'une infinité de couples). Ce fait est évident géométriquement.

De la définition il résulte que si la fonction réciproque est monovalente, c'est-à-dire est une fonction au sens usuel, l'ensemble Y des valeurs de f est l'ensemble de définition de la fonction réciproque φ et l'ensemble de définition X de f, ensemble des valeurs de la fonction réciproque φ. Supposons par exemple qu'une fonction $y = f(x)$ est définie sur un intervalle $[a, b]$, que le segment $[\alpha, \beta]$ est son ensemble des valeurs et que chaque $y \in [\alpha, \beta]$ est associé à un seul $x \in [a, b]$. Alors par définition la fonction réciproque $x = \varphi(y)$ est définie sur le segment $[\alpha, \beta]$ et a pour ensemble des valeurs le segment $[a, b]$ (fig. 61).

La fonction $y = f(x)$ et sa réciproque $x = \varphi(y)$ sont donc représentées par le même graphique. Par exemple, la fonction $y = 5x$ et sa réciproque sont représentées graphiquement par la même droite.

Si l'on permute les axes Ox et Oy, c'est-à-dire si l'on fait subir au plan Oxy une rotation d'angle π autour de la bissectrice du premier quadrant, l'image du graphique de la fonction réciproque $x = \varphi(y)$ sera le graphique de la fonction réciproque $y = \varphi(x)$ (fig. 61).

2. Théorème de continuité de la fonction réciproque.

Théorème 4.15. *Soit $y = f(x)$ une fonction définie, strictement monotone et continue sur un ensemble X et soit Y son ensemble des valeurs. Alors la fonction réciproque $x = \varphi(y)$ est monovalente, strictement monotone et continue sur l'ensemble Y.*

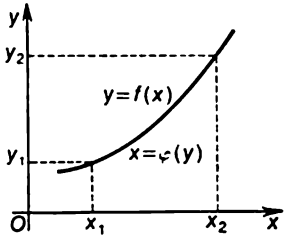


Fig. 62

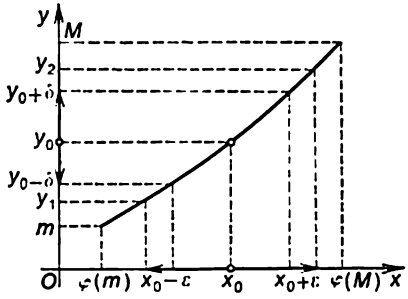


Fig. 63

Démonstration. Supposons pour fixer les idées que la fonction $f(x)$ est strictement croissante sur X (fig. 62).

La fonction réciproque $x = \varphi(y)$ est monovalente, puisque du fait que la fonction $y = f(x)$ est strictement croissante sur X il résulte que $y_1 = f(x_1) \neq f(x_2) = y_2$ pour $x_1 \neq x_2$, donc à chaque $y \in Y$ est associée une seule valeur de $x \in X$.

Montrons maintenant que la fonction réciproque $x = \varphi(y)$ est strictement croissante sur Y. En effet, si $y_1 < y_2$, alors $x_1 < x_2$ ($x_1 = \varphi(y_1)$ et $x_2 = \varphi(y_2)$), puisque si $x_1 \geqslant x_2$ la stricte croissance de $f(x)$ aurait entraîné $y_1 \geqslant y_2$, ce qui contredit l'hypothèse $y_1 < y_2$. Ceci établit donc la stricte monotonie de la fonction réciproque $x = \varphi(y)$.

Montrons enfin que la fonction réciproque $x = \varphi(y)$ est continue sur Y. D'après le corollaire du théorème 4.10, l'ensemble Y est un intervalle de bornes m et M, où $m = \inf\limits_{X} f(x)$ et $M = \sup\limits_{X} f(x)$. Soient $y_0 \in Y$ et $x_0 = \varphi(y_0)$. Traitons d'abord le cas où $m < y_0 < M$ (fig. 63). Le point x_0 est alors manifestement intérieur à l'intervalle X. Prenons $\varepsilon > 0$ tel que $(x_0 - \varepsilon) \in X$ et $(x_0 + \varepsilon) \in X$ et posons $y_1 = f(x_0 - \varepsilon)$ et $y_2 = f(x_0 + \varepsilon)$. La fonction $f(x)$ étant strictement croissante, on a

$$y_1 < y_0 < y_2.$$

Prenons maintenant $\delta > 0$ tel que $y_1 \leqslant y_0 - \delta$ et $y_0 + \delta \leqslant y_2$. Si y vérifie la double inégalité

$$y_0 - \delta < y < y_0 + \delta,$$

alors

$$y_1 < y < y_2,$$

et par suite

$$\varphi(y_1) < \varphi(y) < \varphi(y_2)$$

puisque $\varphi(y)$ est strictement croissante. Vu que $\varphi(y_1) = x_0 - \varepsilon = \varphi(y_0) - \varepsilon$ et $\varphi(y_2) = x_0 + \varepsilon = \varphi(y_0) + \varepsilon$, on obtient $\varphi(y_0) - \varepsilon < \varphi(y) < \varphi(y_0) + \varepsilon$ pourvu que $y_0 - \delta < y < y_0 + \delta$.

On a ainsi prouvé que pour tout $\varepsilon > 0$ assez petit il existe un $\delta > 0$ tel que pour tout y vérifiant l'inégalité $|y - y_0| < \delta$ on

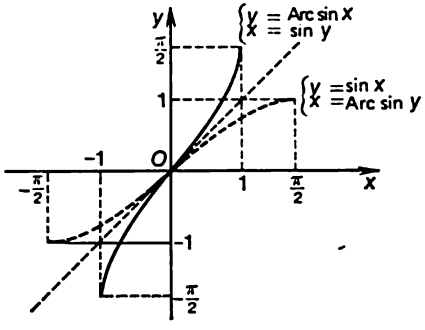


Fig. 64

a $|\varphi(y) - \varphi(y_0)| < \varepsilon$, c'est-à-dire que la fonction réciproque $\varphi(y)$ est continue en y_0. Or y_0 est un point arbitraire de l'intervalle $]m, M[$. Donc la fonction réciproque $\varphi(y)$ est continue sur $]m, M[$ tout entier.

Si $m \in Y$ ou $M \in Y$, on établit par des raisonnements analogues la continuité à droite en m et à gauche en M de la fonction $\varphi(y)$.

Nous avons ainsi prouvé la continuité de la fonction réciproque $x = \varphi(y)$ sur Y.

Le théorème se prouve de façon analogue *mutatis mutandis* si $f(x)$ est strictement décroissante. ■

Remarque. Si la fonction réciproque $x = \varphi(y)$ est monovalente, il est évident que la fonction $y = f(x)$ est réciproque de la fonction $x = \varphi(x)$. De telles fonctions sont dites *mutuellement réciproques.*

Exemple. La fonction $y = \sin x$ est strictement croissante et continue sur le segment $[-\pi/2, \pi/2]$, son ensemble des valeurs est le segment $[-1, 1]$. D'après le théorème 4.15, sa fonction réciproque est définie, continue et strictement croissante sur $[-1, 1]$ et son ensemble des valeurs est le segment $[-\pi/2, \pi/2]$. Cette fonction réciproque est désignée par $x = \text{Arc sin } y$. Elle est représentée par le même graphique que la fonction $y = \sin x$ pour $x \in [-\pi/2, \pi/2]$ (fig. 64).

Si maintenant on permute x et y, c'est-à-dire si l'on envisage la fonction $y = \text{Arc sin } x$, on obtient le graphique tracé par une ligne pleine sur la figure 64.

CHAPITRE V

DÉRIVATION

§ 1. Notion de dérivée

1. Définition de la dérivée. Soit $y = f(x)$ une fonction définie sur un intervalle X. Prenons un point quelconque x_0 de X et donnons à x au point x_0 un accroissement arbitraire Δx tel que le point $x_0 + \Delta x$ appartienne aussi à X. L'accroissement correspondant de la fonction est : $\Delta y = f(x_0 + \Delta x) - f(x_0)$.

Définition. On appelle *dérivée de la fonction* $y = f(x)$ *au point* x_0 la limite (si elle existe) du rapport de l'accroissement de cette fonction à l'accroissement de l'argument lorsque $\Delta x \to 0$.

La dérivée de la fonction $y = f(x)$ en x_0 se désigne par $y'(x_0)$ ou $f'(x_0)$.

Ainsi, par définition,

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Cette dérivée s'appelle *dérivée finie.*

Si

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \pm \infty$$

pour une valeur x_0, on dit que la fonction y admet une *dérivée infinie* au point x_0. Si une fonction $f(x)$ admet une dérivée $f'(x)$ continue en tout point $x \in X$, cette dérivée peut être traitée comme une fonction de x définie aussi sur X.

La définition de la dérivée est constructive.

Exemple. Calculer la dérivée de la fonction $f(x) = x^2$ en $x = x_0$.

S o l u t i o n. Donnons à x un accroissement Δx au point x_0 et cherchons l'accroissement correspondant de la fonction $f(x)$:

$$\Delta y = f(x_0 + \Delta x) - f(x_0) = (x_0 + \Delta x)^2 - x_0^2 =$$
$$= x_0^2 + 2x_0 \Delta x + (\Delta x)^2 - x_0^2 = 2x_0 \Delta x + (\Delta x)^2.$$

Formons le rapport

$$\frac{\Delta y}{\Delta x} = \frac{2x_0 \Delta x + (\Delta x)^2}{\Delta x}.$$

Calculons la limite de ce rapport lorsque $\Delta x \to 0$:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{2x_0 \Delta x + (\Delta x)^2}{\Delta x} = 2x_0.$$

Donc la dérivée de la fonction $f(x) = x^2$ en x_0 est égale à $2x_0$. Ceci se note: $f'(x_0) = 2x_0$.

2. **Signification géométrique de la dérivée.** Soient $y = f(x)$ une fonction définie sur un intervalle $]a, b[$, M, le point du graphique de cette fonction d'abscisse x_0, P, le point d'abscisse $x_0 + \Delta x$. Menons par les points M et P une droite que l'on appellera *sécante*. Désignons par $\varphi(\Delta x)$ l'angle $(\overrightarrow{Ox}, \overrightarrow{MP})$ (fig. 65). Il est évident que cet angle dépend de Δx.

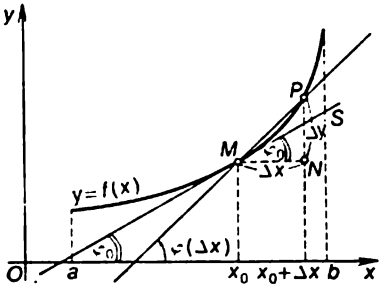


Fig. 65

Si $\lim_{\Delta x \to 0} \varphi(\Delta x) = \varphi_0$, la droite de coefficient directeur $k = \operatorname{tg} \varphi_0$ passant par le point $M(x_0; f(x_0))$ s'appelle *position limite de la sécante MP* lorsque $\Delta x \to 0$ (ou lorsque $P \to M$).

Définition. On appelle *tangente* à la courbe représentative d'une fonction $y = f(x)$ en un point M la position limite de la sécante MP lorsque $\Delta x \to 0$ ou, ce qui est équivalent, lorsque $P \to M$.

De la définition il résulte qu'une condition suffisante d'existence de la tangente est l'existence de $\lim_{\Delta x \to 0} \varphi(\Delta x) = \varphi_0$, où φ_0 est l'angle de la tangente et de l'axe Ox.

Montrons que si la fonction $y = f(x)$ est dérivable en x_0, la courbe représentative de $y = f(x)$ admet une tangente au point $M(x_0; f(x_0))$, et de plus le coefficient directeur de cette tangente est égal à la valeur de la dérivée en x_0, c'est-à-dire à $f'(x_0)$.

En effet, du triangle MNP on déduit que

$$\operatorname{tg} \varphi(\Delta x) = \frac{\Delta y}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

D'où

$$\varphi(\Delta x) = \operatorname{Arc\,tg} \frac{\Delta y}{\Delta x}. \qquad (1)$$

Faisons tendre Δx vers 0 dans l'égalité (1). Puisque $f'(x_0)$ existe, il en est de même de

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x_0).$$

Ceci et la continuité de la fonction Arc tg $\frac{\Delta y}{\Delta x}$ entraînent l'existence de la limite du second membre de l'égalité (1):

$$\lim_{\Delta x \to 0} \operatorname{Arc\,tg} \frac{\Delta y}{\Delta x} = \operatorname{Arc\,tg} \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \operatorname{Arc\,tg} f'(x_0).$$

Donc la limite du premier membre de (1) existe aussi et l'on obtient

$$\lim_{\Delta x \to 0} \varphi(\Delta x) = \operatorname{Arc\,tg} f'(x_0).$$

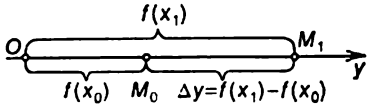


Fig. 66

Or, ceci exprime que la sécante MP possède une position limite, autrement dit la courbe représentative de la fonction $y = f(x)$ admet une tangente en $M(x_0; f(x_0))$ qui fait avec Ox un angle φ_0 égal à Arc tg $f'(x_0)$ et dont le coefficient directeur est donc égal à $\operatorname{tg} \varphi_0 = f'(x_0)$. C.Q.F.D.

Ainsi, la dérivée de la fonction $y = f(x)$ au point x_0 est égale au coefficient directeur de la tangente à la courbe $y = f(x)$ au point $M(x_0; f(x_0))$.

3. Signification mécanique de la dérivée. Supposons que la fonction $y = f(x)$ décrit la loi du mouvement d'un point matériel M sur une droite, autrement dit $y = f(x)$ représente le chemin parcouru par le point M à partir de l'origine au bout d'un temps x.

A la date x_0 le chemin parcouru est $y = f(x_0)$ et à la date x_1, $y = f(x_1)$. Durant l'intervalle de temps $\Delta x = x_1 - x_0$, le point M parcourt le chemin $\Delta y = f(x_1) - f(x_0) = f(x_0 + \Delta x) - f(x_0)$ (fig. 66). Le rapport $\frac{\Delta y}{\Delta x}$ s'appelle *vitesse moyenne* du mouvement (v_m) durant Δx, et la limite du rapport $\frac{\Delta y}{\Delta x}$ lorsque $\Delta x \to 0$, *vitesse instantanée* (v_i) du point M à l'instant x_0.

La notion de vitesse qui a été empruntée à la physique se prête bien à l'étude du comportement d'une fonction quelconque. Quelle que soit la nature de la dépendance exprimée par la fonction $y = f(x)$, le rapport $\frac{\Delta y}{\Delta x}$ est la vitesse moyenne de variation de y par rapport à la variation de x et $y'(x_0)$, la vitesse instantanée de variation de y pour $x = x_0$.

L'importance de la dérivée dans l'étude de tout processus ou phénomène naturel est qu'elle permet d'apprécier la vitesse de variation des variables décrivant ce processus ou phénomène.

4. Dérivées à droite et à gauche. Introduisons les notions de dérivées à droite et à gauche à l'aide de la notion de limite à droite et à gauche d'une fonction.

Définition. On appelle *dérivée à droite* (resp. *à gauche*) d'une fonction $y = f(x)$ en un point x_0 la limite à droite (resp. à gauche), si elle existe, du rapport $\frac{\Delta y}{\Delta x}$ lorsque $\Delta x \to 0$.

Notations:

$$f'_+(x_0) = \lim_{\Delta x \to 0+} \frac{\Delta y}{\Delta x} \left(\text{resp. } f'_-(x_0) = \lim_{\Delta x \to 0-} \frac{\Delta y}{\Delta x}\right).$$

Si une fonction $f(x)$ est dérivable en un point x_0, elle admet en ce point des dérivées à droite et à gauche qui sont confondues.

Il existe cependant des fonctions possédant une dérivée à droite et une dérivée à gauche en un point, mais qui ne sont pas dérivables en ce point. Exemple: la fonction $f(x) = |x|$ dont la dérivée à droite en $x = 0$ est $f'_+(0) = \lim\limits_{\Delta x \to 0+} \frac{\Delta y}{\Delta x} = 1$ ($\Delta y = \Delta x$ pour $x \geqslant 0$) et la dérivée à gauche $f'_-(0) = \lim\limits_{\Delta x \to 0-} \frac{\Delta y}{\Delta x} = -1$ ($\Delta y = -\Delta x$ pour $x < 0$) et qui n'est pas dérivable en ce point puisque $f'_+(0) \neq f'_-(0)$.

§ 2. Notion de dérivabilité d'une fonction

1. Notion de dérivabilité d'une fonction en un point donné.

Définition. On dit qu'une fonction $y = f(x)$ est *dérivable en un point* x_0 si son accroissement Δy en ce point peut être représenté sous la forme

$$\Delta y = A\Delta x + \alpha(\Delta x)\,\Delta x, \tag{1}$$

où A est un nombre indépendant de Δx, $\alpha(\Delta x)$ une fonction de Δx qui est infiniment petite lorsque $\Delta x \to 0$, c'est-à-dire que $\lim\limits_{\Delta x \to 0} \alpha(\Delta x) = 0$.

Établissons un lien entre la dérivabilité d'une fonction en un point et l'existence de la dérivée en ce point.

Théorème 5.1. *Pour qu'une fonction $y = f(x)$ soit dérivable en un point x_0, il est nécessaire et suffisant qu'elle admette une dérivée finie en ce point.*

Démonstration. *Nécessité.* Supposons qu'une fonction $y = f(x)$ est dérivable en un point x_0, c'est-à-dire que son accroissement en ce point peut être mis sous la forme (1): $\Delta y = A\Delta x + \alpha(\Delta x)\,\Delta x$. Une division de cette égalité par Δx (pour $\Delta x \neq 0$) nous donne

$$\frac{\Delta y}{\Delta x} = A + \alpha(\Delta x).$$

En faisant tendre Δx vers 0, on obtient

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} (A + \alpha(\Delta x)) = A.$$

Il s'ensuit que la dérivée existe en x_0 et vaut A : $f'(x_0) = A$.

Suffisance. Supposons qu'existe la dérivée finie $f'(x_0)$, c'est-à-dire que $\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x_0)$. Supposons que $f'(x_0) = A$; la fonction $\alpha(\Delta x) = \frac{\Delta y}{\Delta x} - A$ est alors un infiniment petit lorsque $\Delta x \to 0$ (cf. théorème 4.5). La dernière égalité entraîne

$$\Delta y = A\Delta x + \alpha(\Delta x),$$

où $\lim_{\Delta x \to 0} \alpha(\Delta x) = 0$. Ce qui prouve que la fonction $y = f(x)$ est dérivable en x_0. ■

Donc la dérivabilité et l'existence de la dérivée sont des notions équivalentes pour les fonctions d'une variable.

Remarque. La fonction $\alpha(\Delta x) = \Delta y/\Delta x - A$ introduite dans la démonstration de la suffisance n'est pas définie pour $\Delta x = 0$. Donc l'expression (1) de Δy ne l'est pas non plus pour $\Delta x = 0$. Si l'on définit $\alpha(0)$ de façon arbitraire, l'égalité (1) sera valable pour $\Delta x = 0$ aussi. Dans la suite on aura intérêt à convenir que dans l'expression (1) la fonction $\alpha(\Delta x)$ est définie par continuité pour $\Delta x = 0$, c'est-à-dire que $\alpha(0) = \lim_{\Delta x \to 0} \alpha(\Delta x) = 0$.

2. Lien entre les notions de dérivabilité et de continuité.

Théorème 5.2. *Si une fonction $y = f(x)$ est dérivable en un point donné x_0, elle est continue en x_0.*

Démonstration. La fonction $y = f(x)$ étant dérivable au point x_0, son accroissement en ce point est susceptible de la représentation (1). En faisant tendre Δx vers 0, on obtient alors

$$\lim_{\Delta x \to 0} \Delta y = A \lim_{\Delta x \to 0} \Delta x + \lim_{\Delta x \to 0} \alpha(\Delta x) \lim_{\Delta x \to 0} \Delta x = 0,$$

ce qui exprime que la fonction $y = f(x)$ est continue en x_0 en vertu de la troisième définition de la continuité d'une fonction en un point. ■

Remarque. La réciproque est fausse. Une fonction peut être continue en un point sans y être dérivable.

Exemple : la fonction $f(x) = |x|$ qui est continue en $x = 0$ mais qui, comme prouvé au n° 4 du § 1, n'est pas dérivable en ce point.

Si une fonction $f(x)$ est dérivable en chaque point d'un intervalle, on dira qu'elle l'est sur cet intervalle tout entier.

§ 3. Notion de différentielle

1. Définition et signification géométrique de la différentielle. Soit $y = f(x)$ une fonction dérivable en un point x_0, c'est-à-dire que son accroissement Δy en ce point se représente par la somme

$$\Delta y = A\Delta x + \alpha(\Delta x)\,\Delta x,$$

où $\lim\limits_{\Delta x \to 0} \alpha(\Delta x) = 0$. Le terme $A\Delta x$ est un infiniment petit du même ordre que Δx lorsque $\Delta x \to 0$ (pour $A \neq 0$), de plus il est linéaire par rapport à Δx. Le terme $\alpha(\Delta x)\,\Delta x$ est un infiniment petit d'ordre supérieur à Δx lorsque $\Delta x \to 0$ ($\lim\limits_{\Delta x \to 0} \frac{\alpha(\Delta x)\Delta x}{\Delta x} = 0$).

Donc le premier terme est la partie principale de l'accroissement de la fonction $y = f(x)$ (pour $A \neq 0$).

Définition. On appelle *différentielle* d'une fonction $y = f(x)$ en un point x_0 la partie principale linéaire par rapport à Δx de l'accroissement de cette fonction en x_0:

$$dy = A\Delta x. \tag{1}$$

Si $A = 0$, on a $A\Delta x = 0$ et par suite le terme $A\Delta x$ n'est plus la partie principale de l'accroissement Δy, puisque le terme $\alpha(\Delta x)\,\Delta x$ est généralement différent de zéro. Mais dans ce cas aussi on convient par définition que la différentielle de la fonction en x_0 est égale à $A\Delta x$, c'est-à-dire que $dy = 0$ ici.

En se servant du théorème 5.1, c'est-à-dire en tenant compte du fait que $A = f'(x_0)$, on peut mettre la formule (1) sous la forme

$$dy = f'(x_0)\,\Delta x. \tag{2}$$

Soit $f(x) = x$. D'après la formule (2),

$$dy = dx = (x_0)'\,\Delta x = \left(\lim_{\Delta x \to 0} \frac{(x_0 + \Delta x) - x_0}{\Delta x}\right)\Delta x =$$

$$= \left(\lim_{\Delta x \to 0} \frac{\Delta x}{\Delta x}\right)\Delta x = 1 \cdot \Delta x = \Delta x.$$

Appelons différentielle de la variable indépendante x l'accroissement de cette variable $dx = \Delta x$. La relation (2) devient maintenant

$$dy = f'(x_0)\,dx. \tag{3}$$

On remarquera que la relation (3) nous permet de calculer la dérivée $f'(x_0)$ comme le rapport de la différentielle dy de la fonction y à la différentielle dx de la variable indépendante x, soit

$$f'(x_0) = \frac{dy}{dx}.$$

La différentielle d'une fonction admet l'interprétation géométrique suivante. Soient M et P des points de la courbe représentative d'une fonction $y = f(x)$ d'abscisses respectives x_0 et $x_0 + \Delta x$; MS, la tangente à la courbe $y = f(x)$ en M; α, l'angle de la tangente MS et de l'axe Ox; MN et PN, les parallèles à Ox et Oy respectivement; Q, le point d'intersection de MS et PN (fig. 67). L'accroissement Δy est numériquement égal à la longueur du segment NP. Le triangle rectangle MNQ nous donne: $NQ = \Delta x \operatorname{tg} \alpha = f'(x_0)\,\Delta x = \mathrm{d}y$, c'est-à-dire que la différentielle de la fonction $\mathrm{d}y$ est égale à la longueur du segment NQ. Il est évident que $NP \neq NQ$.

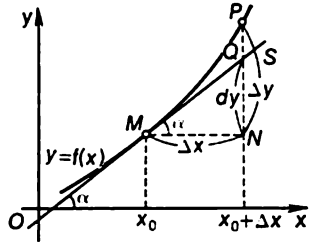


Fig. 67

Donc la différentielle $\mathrm{d}y$ de la fonction $y = f(x)$ au point x_0 est égale à l'accroissement de l'ordonnée de la tangente MS au graphique de cette fonction en $M(x_0; f(x_0))$, quant à l'accroissement Δy, il est égal à celui de l'ordonnée de la fonction $y = f(x)$ en x_0 correspondant à Δx.

2. Calculs approchés à l'aide de la différentielle. La différentielle, de par sa définition, dépend linéairement de Δx et est la partie principale de l'accroissement Δy de la fonction y. L'accroissement Δy dépend de façon plus complexe de Δx. Si, par exemple, $f(x) = x^3$, on a $\Delta y = (x_0 + \Delta x)^3 - x_0^3 = 3x_0^2\,\Delta x + 3x_0\,(\Delta x)^2 + (\Delta x)^3$, alors que

$$\mathrm{d}y = f'(x_0)\,\Delta x = \left(\lim_{\Delta x \to 0} \frac{(x_0+\Delta x)^3 - x_0^3}{\Delta x}\right)\Delta x = 3x_0^2\,\Delta x.$$

Dans de nombreux problèmes l'accroissement d'une fonction en un point est remplacé approximativement par la différentielle de cette fonction en ce point:

$$\Delta y \approx \mathrm{d}y.$$

L'erreur absolue de cette approximation est égale à $|\,\Delta y - \mathrm{d}y\,|$ et est un infiniment petit d'ordre supérieur à Δx lorsque $\Delta x \to 0$.

Exemple. Montrons que pour α petit on peut se servir de la formule approchée

$$\sqrt{1+\alpha} \approx 1 + \alpha/2.$$

Solution. Considérons la fonction $f(x) = \sqrt{x}$. Pour Δx petit on a

$$\Delta y = \sqrt{x_0+\Delta x} - \sqrt{x_0} \approx dy \quad \text{ou} \quad \sqrt{x_0+\Delta x} - \sqrt{x_0} \approx$$

$$\approx (\sqrt{x})'|_{x=x_0}\Delta x = \left(\lim_{\Delta x\to 0} \frac{\sqrt{x_0+\Delta x}-\sqrt{x_0}}{\Delta x}\right)\Delta x =$$

$$= \left(\lim_{\Delta x\to 0} \frac{(x_0+\Delta x)-x_0}{\Delta x\,(\sqrt{x_0+\Delta x}+\sqrt{x_0})}\right)\Delta x = \frac{1}{2\sqrt{x_0}}\Delta x,$$

d'où, en posant $x_0 = 1$, $\Delta x = \alpha$,

$$\sqrt{1+\alpha} \approx 1+\alpha/2.$$

En particulier, $\sqrt{1{,}0003} \approx 1{,}00015$ pour $\alpha = 0{,}0003$.

Établissons maintenant les règles de dérivation et de calcul des dérivées des fonctions élémentaires. Remarquons seulement que dans ces formules on écrit généralement x au lieu de x_0 et on admet que x est fixe.

§ 4. Règles de dérivation d'une somme, d'une différence, d'un produit et d'un quotient

Théorème 5.3. *Si des fonctions $u = u(x)$ et $v = v(x)$ sont dérivables en un point x, il en est de même de leur somme, leur différence, leur produit et leur quotient (sous réserve que $v(x) \neq 0$) et de plus*

$$(u \pm v)' = u' \pm v', \quad (u\cdot v)' = u'v + uv', \quad \left(\frac{u}{v}\right)' = \frac{u'v-uv'}{v^2}. \qquad (1)$$

Démonstration. Pour établir les formules (1), on se servira de la définition de la dérivée, de l'égalité $f(x+\Delta x) = f(x) + \Delta y$ et du théorème 4.3. On a alors

$$(u \pm v)' = \lim_{\Delta x\to 0} \frac{[u(x+\Delta x) \pm v(x+\Delta x)] - [u(x) \pm v(x)]}{\Delta x} =$$

$$= \lim_{\Delta x\to 0} \left[\frac{u(x+\Delta x)-u(x)}{\Delta x} \pm \frac{v(x+\Delta x)-v(x)}{\Delta x}\right] =$$

$$= \lim_{\Delta x\to 0} \frac{u(x+\Delta x)-u(x)}{\Delta x} \pm \lim_{\Delta x\to 0} \frac{v(x+\Delta x)-v(x)}{\Delta x} =$$

$$= \lim_{\Delta x\to 0} \frac{\Delta u}{\Delta x} \pm \lim_{\Delta x\to 0} \frac{\Delta v}{\Delta x} = u' \pm v';$$

$$(uv)' = \lim_{\Delta x\to 0} \frac{u(x+\Delta x)\,v(x+\Delta x) - u(x)\,v(x)}{\Delta x} =$$

$$= \lim_{\Delta x\to 0} \frac{[u(x)+\Delta u]\,[v(x)+\Delta v] - u(x)\,v(x)}{\Delta x} =$$

$$= \lim_{\Delta x \to 0} \frac{u(x)\,v(x) + \Delta u v(x) + u(x)\,\Delta v + \Delta u\,\Delta v - u(x)\,v(x)}{\Delta x} =$$

$$= \lim_{\Delta x \to 0} \left[v(x) \frac{\Delta u}{\Delta x} + u(x) \frac{\Delta v}{\Delta x} + \Delta v \frac{\Delta u}{\Delta x} \right] =$$

$$= v \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} + u \lim_{\Delta x \to 0} \frac{\Delta v}{\Delta x} + \lim_{\Delta x \to 0} \Delta v \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} =$$

$$= vu' + uv' + 0 \cdot u' = u'v + uv',$$

puisque $\lim\limits_{\Delta x \to 0} \Delta v = 0$ et les facteurs u et v sont indépendants de Δx;

$$\left(\frac{u}{v}\right)' = \lim_{\Delta x \to 0} \frac{\frac{u(x+\Delta x)}{v(x+\Delta x)} - \frac{u(x)}{v(x)}}{\Delta x} =$$

$$= \lim_{\Delta x \to 0} \frac{u(x+\Delta x)\,v(x) - u(x)\,v(x+\Delta x)}{\Delta x \cdot v(x+\Delta x)\,v(x)} =$$

$$= \lim_{\Delta x \to 0} \frac{[u(x)+\Delta u]\,v(x) - u(x)\,[v(x)+\Delta v]}{\Delta x v(x)\,[v(x)+\Delta v]} =$$

$$= \lim_{\Delta x \to 0} \frac{uv + \Delta uv - uv - u\,\Delta v}{\Delta x v(v+\Delta v)} = \lim_{\Delta x \to 0} \frac{v \frac{\Delta u}{\Delta x} - u \frac{\Delta v}{\Delta x}}{v^2 + v\,\Delta v} =$$

$$= \frac{v \lim\limits_{\Delta x \to 0} \frac{\Delta u}{\Delta x} - u \lim\limits_{\Delta x \to 0} \frac{\Delta v}{\Delta x}}{v^2 + v \lim\limits_{\Delta x \to 0} \Delta v} = \frac{u'v - uv'}{v^2}. \quad \blacksquare$$

§ 5. Calcul des dérivées des fonctions constante, puissance, circulaires et logarithme

1. Dérivée d'une fonction constante. La dérivée de la fonction $y = f(x) = C$, où C est une constante, s'exprime par la formule

$$y' = 0.$$

Démonstration. Quels que soient x et Δx, on a $f(x + \Delta x) = C$ et $\Delta y = f(x + \Delta x) - f(x) = 0$. D'où $\frac{\Delta y}{\Delta x} = 0$ pour tout $\Delta x \neq 0$ et par suite

$$y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = 0. \quad \blacksquare$$

2. Dérivée de la fonction puissance. La dérivée de la fonction $y = x^n$, où l'exposant n est un entier >0, est donnée par la formule

$$y' = n \cdot x^{n-1}.$$

Démonstration. Grâce à la formule du binôme de Newton, on peut écrire

$$\Delta y = (x+\Delta x)^n - x^n =$$

$$= \left[x^n + nx^{n-1}\Delta x + \frac{n(n-1)}{2!}x^{n-2}(\Delta x)^2 + \ldots + (\Delta x)^n\right] - x^n =$$

$$= nx^{n-1}\Delta x + \frac{n(n-1)}{2!}x^{n-2}(\Delta x)^2 + \ldots + (\Delta x)^n.$$

Donc pour $\Delta x \neq 0$ on a

$$\frac{\Delta y}{\Delta x} = nx^{n-1} + \frac{n(n+1)}{2!}x^{n-2}\Delta x + \ldots + (\Delta x)^{n-1}.$$

Puisque $\lim\limits_{\Delta x\to 0}\Delta x = 0$, $\lim\limits_{\Delta x\to 0}(\Delta x)^2 = 0$, ..., $\lim\limits_{\Delta x\to 0}(\Delta x)^{n-1} = 0$, il vient

$$y' = \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = nx^{n-1}. \quad \blacksquare$$

Remarque. Le cas d'une fonction puissance dont l'exposant est un réel quelconque est traité au n° 2 du § 9.

3. Dérivées des fonctions circulaires.

1) La dérivée de la fonction $y = \sin x$ est

$$y' = \cos x.$$

Démonstration. On a

$$\Delta y = \sin(x+\Delta x) - \sin x =$$

$$= 2\sin(\Delta x/2)\cos(x+\Delta x/2).$$

Donc, pour $\Delta x \neq 0$

$$\frac{\Delta y}{\Delta x} = \frac{2\sin(\Delta x/2)\cos(x+\Delta x/2)}{\Delta x} = \frac{\sin(\Delta x/2)}{\Delta x/2}\cos(x+\Delta x/2).$$

Comme $\lim\limits_{\Delta x\to 0}\dfrac{\sin(\Delta x/2)}{\Delta x/2} = 1$ (la première limite remarquable) et que $\lim\limits_{\Delta x\to 0}\cos(x+\Delta x/2) = \cos x$ en vertu de la continuité de $\cos x$, il vient

$$y' = \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \cos x. \quad \blacksquare$$

2) La dérivée de la fonction $y = \cos x$ est

$$y' = -\sin x.$$

Démonstration. On a

$$\Delta y = \cos(x+\Delta x) - \cos x = -2\sin(\Delta x/2)\sin(x+\Delta x/2).$$

Donc, pour $\Delta x \neq 0$

$$\frac{\Delta y}{\Delta x} = -\frac{2 \sin(\Delta x/2) \sin(x+\Delta x/2)}{\Delta x} = -\frac{\sin(\Delta x/2)}{\Delta x/2} \sin(x+\Delta x/2).$$

Puisque $\lim\limits_{\Delta x \to 0} \sin\left(x + \frac{\Delta x}{2}\right) = \sin x$ en raison de la continuité de $\sin x$, il vient

$$y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\sin x. \quad \blacksquare$$

3) La dérivée de la fonction $y = \operatorname{tg} x$ est

$$y' = \frac{1}{\cos^2 x} \left(x \neq \frac{\pi}{2} + n\pi\right).$$

Démonstration. Comme $\operatorname{tg} x = \sin x/\cos x$, le théorème 5.3 nous donne

$$y' = \frac{(\sin x)' \cos x - \sin x (\cos x)'}{\cos^2 x} =$$

$$= \frac{\cos x \cos x - \sin x (-\sin x)}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x},$$

et par suite

$$y' = \frac{1}{\cos^2 x}.$$

4) La dérivée de la fonction $y = \operatorname{cotg} x$ est

$$y' = -\frac{1}{\sin^2 x} \quad (x \neq n\pi).$$

Démonstration. Comme $\operatorname{cotg} x = \cos x/\sin x$, on a par analogie à ce qui précède

$$y' = \frac{(\cos x)' \sin x - \cos x (\sin x)'}{\sin^2 x} =$$

$$= \frac{(-\sin x) \sin x - \cos x \cos x}{\sin^2 x} = -\frac{\sin^2 x + \cos^2 x}{\sin^2 x},$$

donc

$$y' = -\frac{1}{\sin^2 x}. \quad \blacksquare$$

4. Dérivée de la fonction logarithme. La dérivée de la fonction $y = \log_a x$ $(0 < a \neq 1)$ est égale à

$$y' = \frac{1}{x} \log_a e = \frac{1}{x \ln a}.$$

Démonstration. On a

$$\Delta y = \log_a (x+\Delta x) - \log_a x = \log_a \frac{x+\Delta x}{x} = \log_a \left(1 + \frac{\Delta x}{x}\right).$$

Donc, pour $\Delta x \neq 0$

$$\frac{\Delta y}{\Delta x} = \frac{1}{\Delta x} \log_a \left(1 + \frac{\Delta x}{x}\right) = \frac{1}{x} \frac{x}{\Delta x} \log_a \left(1 + \frac{\Delta x}{x}\right),$$

ou

$$\frac{\Delta y}{\Delta x}=\frac{1}{x}\log_a\left[\left(1+\frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}\right].$$

En posant $x/\Delta x = h$, on obtient $\lim\limits_{\Delta x\to 0}(1+\Delta x/x)^{x/\Delta x} = \lim\limits_{h\to\infty}(1+1/h)^h = e$ · (la deuxième limite remarquable) et puisque la fonction logarithme est continue, il vient

$$y'=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\frac{1}{x}\log_a\left[\lim_{\Delta x\to 0}\left(1+\frac{\Delta x}{x}\right)^{x/\Delta x}\right]=$$

$$=\frac{1}{x}\log_a e=\frac{1}{x\ln a}.\ \blacksquare$$

Corollaire. *Si* $y=\log_e x=\ln x$, *alors* $y'=(\ln x)'=1/x$.

§ 6. Théorème de la dérivée de la fonction réciproque

Soient $y=f(x)$ une fonction satisfaisant les conditions du théorème 4.15 de la fonction réciproque et $x=\varphi(y)$ sa réciproque. On a alors le théorème suivant.

Théorème 5.4. *Si la fonction* $y=f(x)$ *admet en un point* x_0 *une dérivée* $f'(x_0)$ *non nulle, sa réciproque* $x=\varphi(y)$ *admet aussi une dérivée non nulle* $\varphi'(y_0)$ *au point* $y_0=f(x_0)$, *et de plus*

$$\varphi'(y_0)=\frac{1}{f'(x_0)}.$$

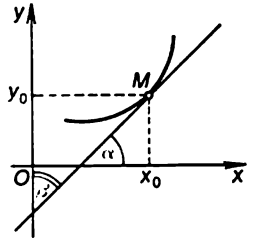


Fig. 68

Démonstration. Donnons à l'argument y de la fonction réciproque $x=\varphi(y)$ un accroissement $\Delta y\neq 0$ au point y_0. La fonction $x=\varphi(y)$ subit un accroissement Δx qui est $\neq 0$ en vertu de la stricte monotonie de la fonction réciproque. On peut donc écrire

$$\frac{\Delta x}{\Delta y}=\frac{1}{\Delta y/\Delta x}.$$

Faisons tendre Δy vers 0. Puisque la fonction réciproque $x=\varphi(y)$ est continue en y_0 (cf. théorème 4.15), $\Delta x\to 0$ avec Δy. Or, lorsque $\Delta x\to 0$, le second membre admet une limite égale à $1/f'(x_0)$. Donc le premier membre en admet aussi une qui est égale par définition à $\varphi'(y_0)$. On obtient en définitive

$$\varphi'(y_0)=\frac{1}{f'(x_0)}.\ \blacksquare \qquad (1)$$

Le théorème établi admet une interprétation géométrique simple. Considérons le graphique de la fonction $y = f(x)$ (ou celui de sa réciproque $x = \varphi(y)$) au voisinage d'un point x_0. Supposons que x_0 est l'abscisse d'un point M de ce graphique (fig. 68). On sait que la dérivée $f'(x_0)$ est le coefficient directeur de la tangente en M à la courbe $y = f(x)$. La dérivée $\varphi'(y_0)$ de la fonction réciproque est égale au coefficient directeur de cette même tangente (mais par rapport à Oy). Les angles α et β étant complémentaires, la formule (1) exprime le fait évident

$$\varphi'(y_0) = \operatorname{tg}\beta = \frac{1}{\operatorname{cotg}\beta} = \frac{1}{\operatorname{cotg}(\pi/2 - \alpha)} = \frac{1}{\operatorname{tg}\alpha} = \frac{1}{f'(x_0)}.$$

§ 7. Calcul des dérivées de la fonction exponentielle et des fonctions circulaires réciproques

Poursuivons le calcul des dérivées des fonctions élémentaires à l'aide du théorème 5.4 établi plus haut.

1. Dérivée de la fonction exponentielle. La dérivée de la fonction $y = a^x$ $(0 < a \neq 1)$ est égale à

$$y' = a^x \ln a.$$

Démonstration. La fonction exponentielle $y = a^x$ est réciproque de la fonction logarithme $x = \log_a y$. Puisque

$$x'(y) = \frac{1}{y}\log_a e,$$

le théorème 5.4 de la dérivée de la fonction réciproque et la relation classique $\log_a b = 1/\log_b a$ nous donnent

$$y'(x) = \frac{1}{x'(y)} = \frac{y}{\log_a e} = a^x \ln a. \blacksquare$$

Corollaire. *Si* $y = e^x$, *alors* $y' = (e^x)' = e^x$.

2. Dérivées des fonctions circulaires réciproques.
1) La dérivée de la fonction $y = \operatorname{Arc}\sin x$ est

$$y' = \frac{1}{\sqrt{1-x^2}} \quad (|x| < 1).$$

Démonstration. La fonction $y = \operatorname{Arc}\sin x$ est réciproque de la fonction $x = \sin y$. Comme $x'(y) = \cos y$, on a d'après le théorème 5.4 de la dérivée de la fonction réciproque

$$y'(x) = \frac{1}{x'(y)} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-\sin^2 y}}.$$

La racine est prise avec le signe $+$, puisque $\cos y > 0$ lorsque $y \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. Comme $\sin y = x$, on obtient en définitive

$$y'(x) = \frac{1}{\sqrt{1-x^2}}. \blacksquare$$

2) La dérivée de la fonction $y = \operatorname{Arc}\cos x$ est

$$y'(x) = -\frac{1}{\sqrt{1-x^2}}.$$

La démonstration est identique à la précédente.

3) La dérivée de la fonction $y = \operatorname{Arc}\operatorname{tg} x$ est

$$y' = \frac{1}{1+x^2}.$$

Démonstration. La fonction $y = \operatorname{Arc}\operatorname{tg} x$ est réciproque de la fonction $x = \operatorname{tg} y$.

Comme $x'(y) = 1/\cos^2 y$, on a $y'(x) = 1/x'(y) = \cos^2 y$. Or $1/\cos^2 y = 1 + \operatorname{tg}^2 y = 1 + x^2$, donc

$$y'(x) = \frac{1}{1+x^2}. \blacksquare$$

4) La dérivée de la fonction $y = \operatorname{Arc}\operatorname{cotg} x$ est

$$y' = -\frac{1}{1+x^2}.$$

La démonstration est identique à la précédente.

§ 8. Règle de dérivation d'une fonction composée

Théorème 5.5. *Si une fonction $x = \varphi(t)$ est dérivable en un point t_0 et une fonction $y = f(x)$, dérivable au point correspondant $x_0 = \varphi(t_0)$, la fonction composée $f[\varphi(t)]$ est dérivable aussi en t_0 et l'on a la formule suivante*

$$y'(t_0) = f'(x_0)\,\varphi'(t_0). \tag{1}$$

Démonstration. La fonction $y = f(x)$ étant dérivable en x_0, son accroissement en ce point peut être écrit sous la forme

$$\Delta y = f'(x_0)\,\Delta x + \alpha(\Delta x)\,\Delta x, \tag{2}$$

où $\lim\limits_{\Delta x \to 0} \alpha(\Delta x) = 0$. En divisant l'égalité (2) par Δt, on obtient

$$\frac{\Delta y}{\Delta t} = f'(x_0)\frac{\Delta x}{\Delta t} + \alpha(\Delta x)\frac{\Delta x}{\Delta t}. \tag{3}$$

L'égalité (3) est valable pour tout Δx assez petit. Prenons Δx égal à l'accroissement de la fonction $x = \varphi(t)$ correspondant à l'accrois-

sement Δt de t en t_0 et faisons tendre Δt vers 0 dans cette égalité. Etant dérivable en t_0 par hypothèse, la fonction $x = \varphi(t)$ est continue en ce point. Donc, d'après la troisième définition de la continuité d'une fonction en un point, $\Delta x \to 0$ avec Δt. Alors $\alpha(\Delta x)$ tend aussi vers 0, c'est-à-dire qu'on a

$$\lim_{\Delta t \to 0} \left(\alpha(\Delta x) \frac{\Delta x}{\Delta t} \right) = \lim_{\Delta t \to 0} \alpha(\Delta x) \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = 0 \cdot \varphi'(t_0) = 0. \quad (4)$$

D'après la relation (4) le second membre de l'égalité (3) tend vers $f'(x_0)\,\varphi'(t_0)$ lorsque Δt tend vers 0. Donc le premier membre de (3) admet aussi une limite qui, par définition de la dérivée, est égale à la dérivée de la fonction composée $y = f[\varphi(t)]$ au point t_0. Nous avons ainsi prouvé la dérivabilité de la fonction composée et établi la formule (1). ■

Remarque 1. La fonction envisagée dans ce théorème est une fonction composée dans laquelle y dépend de t au moyen d'une variable intermédiaire x. On peut avoir affaire à une dépendance plus complexe — à deux, trois ou plusieurs variables intermédiaires — mais la règle de dérivation reste la même.

Si, par exemple, $y = f(x)$, $x = \varphi(u)$, $u = \psi(v)$ et $v = \chi(t)$, la dérivée $y'(t)$ se calcule à l'aide de la formule

$$y'(t) = f'(x)\,\varphi'(u)\,\psi'(v)\,\chi'(t). \quad (5)$$

Exemple 1. Calculer la dérivée de la fonction $y = e^{\operatorname{Arc\,tg}\, x}$.

Solution. Cette fonction peut être mise sous la forme $y = e^u$, où $u = \operatorname{Arc\,tg}\, x$. D'après la formule (1)

$$y'(x) = y'(u)\,u'(x) = e^u \frac{1}{1+x^2}.$$

En remplaçant u par $\operatorname{Arc\,tg}\, x$, on obtient en définitive

$$y' = e^{\operatorname{Arc\,tg}\, x} \frac{1}{1+x^2}.$$

Exemple 2. Calculer la dérivée de la fonction $y = \operatorname{tg}^2 \sqrt{x^2+1}$.

Solution. Cette fonction peut être mise sous la forme $y = u^2$, où $u = \operatorname{tg} v$, $v = \sqrt{w}$ et $w = x^2 + 1$. La formule (5) nous donne

$$y'(x) = y'(u)\,u'(v)\,v'(w)\,w'(x) = (u^2)'(\operatorname{tg} v)'\left(\sqrt{w}\right)'(x^2+1)' =$$

$$= 2u \sec^2 v \frac{1}{2\sqrt{w}} 2x = \frac{2x \operatorname{tg} \sqrt{x^2+1} \sec^2 \sqrt{x^2+1}}{\sqrt{x^2+1}}.$$

Remarque 2. La dérivée se calcule parfois directement à partir de sa définition. Calculons, par exemple, la dérivée de la fonction

$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{pour } x \neq 0, \\ 0 & \text{pour } x = 0. \end{cases}$$

Pour $x \neq 0$ la dérivée se calcule à l'aide des formules et des règles de dérivation :

$$f'(x) = \left(x^2 \sin \frac{1}{x}\right)' = (x^2)' \sin \frac{1}{x} + \\ + x^2 \left(\sin \frac{1}{x}\right)' \left(\frac{1}{x}\right)' = 2x \sin \frac{1}{x} - \cos \frac{1}{x}.$$

On ne peut se servir de cette expression pour $x = 0$. En ce point on peut calculer la dérivée à l'aide de sa définition :

$$f'(0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \Delta x \sin \frac{1}{\Delta x} = 0$$

(le produit d'un infiniment petit par une fonction bornée est un infiniment petit). Donc

$$f'(x) = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x} & \text{pour } x \neq 0, \\ 0 & \text{pour } x = 0. \end{cases}$$

§ 9. Dérivée logarithmique. Dérivée d'une fonction puissance à exposant réel quelconque. Table des dérivées des fonctions élémentaires

1. Notion de dérivée logarithmique d'une fonction. Calculons la dérivée de la fonction $y = \ln |x|$ $(x \neq 0)$. Puisque $(\ln x)' = 1/x$ et $(\ln(-x))' = (-x)'/-x = 1/x$ (cette égalité a été acquise à l'aide de la règle de dérivation d'une fonction composée), la dérivée de cette fonction est donnée par la formule :

$$y' = (\ln |x|)' = \frac{1}{x}. \tag{1}$$

Appliquons la formule (1) au calcul de la dérivée de la fonction composée $y = \ln |u|$, où $u = f(x)$ est une fonction dérivable. On a

$$y' = (\ln |u|)' = \frac{1}{u} u' = \frac{f'(x)}{f(x)}$$

ou

$$(\ln |f(x)|)' = \frac{f'(x)}{f(x)}. \tag{2}$$

La dérivée $(\ln |f(x)|)'$ s'appelle *dérivée logarithmique* de la fonction $f(x)$. Pour simplifier l'écriture dans le calcul de la dérivée logarithmique, on omet généralement le module de la fonction $f(x)$.

Calculons la dérivée de la fonction exponentielle-puissance $y = u(x)^{v(x)}$, où u et v sont des fonctions de x ($u > 0$) dérivables en x. Puisque $\ln y = v(x) \ln u(x)$, la formule (2) nous donne

$$\frac{y'}{y} = [v(x) \ln u(x)]' = v'(x) \ln u(x) + v(x) \frac{u'(x)}{u(x)}.$$

D'où l'on déduit la formule de la dérivée

$$y' = u(x)^{v(x)} \left[v'(x) \ln u(x) + v(x) \frac{u'(x)}{u(x)} \right]. \quad (3)$$

Exemple. Calculer la dérivée de la fonction $y = x^x$.

Solution. Cette fonction peut être écrite sous la forme $y = u(x)^{v(x)}$ où $u(x) = x$ et $v(x) = x$. La formule (3) entraîne

$$y' = x^x \left[1 \ln x + x \frac{1}{x} \right] = x^x (\ln x + 1).$$

La dérivée de la fonction $y = u(x)^{v(x)}$ peut être calculée par un autre procédé. Mettons cette fonction sous la forme $y = e^{v(x) \ln u(x)}$ et calculons y' :

$$y' = [e^{v(x) \ln u(x)}]' = e^{v(x) \ln u(x)} [v(x) \ln u(x)]' =$$

$$= y \left[v'(x) \ln u(x) + v(x) \frac{u'(x)}{u(x)} \right];$$

en substituant $y = u(x)^{v(x)}$, on obtient de nouveau la formule (3).

La dérivée logarithmique se prête bien au calcul de la dérivée d'une fonction puissance à exposant réel quelconque.

2. Dérivée d'une fonction puissance à exposant réel quelconque. La dérivée de la fonction $y = x^\alpha$ (où α est un réel quelconque) s'exprime par la formule

$$y' = \alpha x^{\alpha - 1}.$$

Démonstration. Comme $y = x^\alpha$, on a $\ln y = \alpha \ln x$. La formule (2) nous donne

$$\frac{y'}{y} = [\alpha \ln x]' = \frac{\alpha}{x}.$$

D'où, puisque $y = x^\alpha$,

$$y' = (x^\alpha)' = \alpha x^{\alpha - 1}. \blacksquare$$

Récapitulons les calculs des dérivées des fonctions élémentaires en dressant la table suivante.

3. Table des dérivées des fonctions élémentaires.

I. $(C)' = 0$.

II. $(x^\alpha)' = \alpha x^{\alpha-1}$, en particulier $\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$, $(\sqrt{x})' = \frac{1}{2\sqrt{x}}$.

III. $(\log_a x)' = \frac{1}{x}\log_a e$, en particulier $(\ln x)' = \frac{1}{x}$.

IV. $(a^x)' = a^x \ln a$, en particulier, $(e^x)' = e^x$.

V. $(\sin x)' = \cos x$.

VI. $(\cos x)' = -\sin x$.

VII. $(\operatorname{tg} x)' = \frac{1}{\cos^2 x}$.

VIII. $(\operatorname{cotg} x)' = -\frac{1}{\sin^2 x}$.

IX. $(\operatorname{Arc}\sin x)' = \frac{1}{\sqrt{1-x^2}}$.

X. $(\operatorname{Arc}\cos x)' = -\frac{1}{\sqrt{1-x^2}}$.

XI. $(\operatorname{Arc}\operatorname{tg} x)' = \frac{1}{1+x^2}$.

XII. $(\operatorname{Arc}\operatorname{cotg} x)' = -\frac{1}{1+x^2}$.

Les formules de cette table et les règles de dérivation de la somme, du produit, de la différence et du quotient de fonctions ainsi que la règle de dérivation d'une fonction composée sont les principales formules du calcul différentiel. Ces règles et formules nous permettent de tirer une conclusion importante : la dérivée de toute fonction élémentaire est une fonction élémentaire. Donc la dérivation est une opération stable pour la classe des fonctions élémentaires.

§ 10. Dérivées et différentielles supérieures

1. Notion de dérivées d'ordre *n*. Comme déjà mentionné au § 1 de ce chapitre, la dérivée $f'(x)$ d'une fonction $y = f(x)$ est elle-même une fonction de x. On peut donc parler de l'existence et du calcul de sa dérivée.

Appelons $f'(x)$ *dérivée d'ordre un* (ou *première*) de la fonction $f(x)$.

La dérivée de la dérivée d'une fonction s'appelle *dérivée d'ordre deux* (ou *dérivée seconde*) de cette fonction. La dérivée de la dérivée seconde s'appelle *dérivée d'ordre trois* (ou *dérivée troisième*) et ainsi de suite. Les dérivées d'ordre supérieur à deux s'appellent *dérivées supérieures* et se notent y'', y''', $y^{(4)}$, ..., $y^{(n)}$, ... (au lieu de y'' et y''' on écrit parfois $y^{(2)}$ et $y^{(3)}$) ou $f''(x)$, $f'''(x)$, $f^{(4)}(x)$, ..., $f^{(n)}(x)$, ...

La dérivée d'ordre n est la dérivée de la dérivée d'ordre $n-1$, c'est-à-dire que $y^{(n)} = (y^{(n-1)})'$.

Les dérivées supérieures sont largement utilisées en physique. Limitons-nous à l'interprétation physique de la dérivée seconde $f''(x)$. Si une fonction $y = f(x)$ décrit le mouvement d'un point matériel sur une ligne droite, la dérivée première $f'(x)$ est la vitesse instantanée à l'instant x, la dérivée seconde, la vitesse de variation de la vitesse, c'est-à-dire l'accélération du point mobile à cet instant.

2. Formules des dérivées d'ordre n de certaines fonctions.

1) Calculons la dérivée d'ordre n de la fonction puissance $y = x^\alpha$ $(x > 0)$ (α est un réel quelconque). Une dérivation successive nous donne *):

$$y' = \alpha x^{\alpha-1}, \quad y^{(2)} = \alpha(\alpha-1)x^{\alpha-2},$$

$$y^{(3)} = \alpha(\alpha-1)(\alpha-2)x^{\alpha-3}, \ldots,$$

$$y^{(n)} = \alpha(\alpha-1)(\alpha-2)\ldots[\alpha-(n-1)]x^{\alpha-n}.$$

Dans le cas particulier où $\alpha = m$, m étant un entier naturel, on obtient

$$(x^m)^{(m)} = m(m-1)(m-2)\ldots[m-(m-1)]\cdot 1 = m!,$$

$$(x^m)^{(n)} = 0 \text{ pour } n > m.$$

2) Calculons la dérivée d'ordre n de la fonction exponentielle $y = a^x$ $(0 < a \neq 1)$. Une dérivation successive nous donne

$$y' = a^x \ln a, \quad y^{(2)} = a^x(\ln a)^2, \quad y^{(3)} = a^x(\ln a)^3, \quad \ldots$$
$$\ldots y^{(n)} = a^x(\ln a)^n.$$

En particulier, si $y = e^x$, on a pour tout n

$$(e^x)^{(n)} = e^x.$$

3) Calculons la dérivée d'ordre n de la fonction $y = \sin x$. Une dérivation successive nous donne

$$y' = \cos x = \sin\left(x + \frac{\pi}{2}\right),$$

$$y^{(2)} = -\sin x = \sin(x+\pi) = \sin\left(x + 2\frac{\pi}{2}\right),$$

$$y^{(3)} = -\cos x = \sin\left(x + 3\frac{\pi}{2}\right), \ldots, \quad y^{(n)} = \sin\left(x + n\frac{\pi}{2}\right).$$

Donc la dérivée d'ordre n quelconque de $\sin x$ se calcule à

*) Les formules des dérivées d'ordre n s'établissent de façon rigoureuse par récurrence.

l'aide de la formule

$$(\sin x)^{(n)} = \sin\left(x + n\frac{\pi}{2}\right).$$

Par exemple, $(\sin x)^{(10)} = \sin(x + 10(\pi/2)) = \sin(x + \pi) = -\sin x$.

4) On calcule de façon analogue la dérivée d'ordre n de la fonction $y = \cos x$:

$$(\cos x)^{(n)} = \cos\left(x + n\frac{\pi}{2}\right).$$

3. Formule de Leibniz pour la dérivée d'ordre n du produit de deux fonctions. Soit $y = u \cdot v$, où u et v sont des fonctions de x dérivables autant de fois qu'on le veut. Alors

$$y' = u'v + uv', \quad y'' = u''v + u'v' + u'v' + uv'' =$$

$$= u''v + 2u'v' + uv'' = u^{(2)}v + 2u'v' + uv^{(2)},$$

$$y''' = u'''v + u''v' + 2u''v' + 2u'v'' + u'v'' + uv''' =$$

$$= u'''v + 3u''v' + 3u'v'' + uv''' =$$

$$= u^{(3)}v + 3u^{(2)}v' + 3u'v^{(2)} + uv^{(3)}.$$

Les seconds membres de ces égalités ressemblent aux développements des diverses puissances du binôme $(u + v)^n$ à l'aide de la formule de Newton, mais au lieu des exposants on a des nombres définissant l'ordre des dérivées; quant aux fonctions u et v, il faut les traiter, pour une complète analogie avec la formule de Newton, comme des « dérivées d'ordre 0 »: $u^{(0)}$ et $v^{(0)}$. La formule de la dérivée d'ordre n du produit de deux fonctions est donc:

$$y^{(n)} = (uv)^{(n)} = u^{(n)}v + nu^{(n-1)}v' + \frac{n(n-1)}{2!}u^{(n-2)}v'' + \ldots$$

$$\ldots + \frac{n(n-1)\ldots(n-k+1)}{k!}u^{(n-k)}v^{(k)} + \ldots + uv^{(n)}. \qquad (1)$$

La formule (1) s'appelle *formule de Leibniz* *). Prouvons-la par récurrence.

Pour $n = 1$, cette formule s'écrit $(uv)' = u'v + uv'$, ce qui est confondu avec la formule de dérivation du produit de deux fonctions. Pour $n = 2$ et $n = 3$ elle est également vérifiée. Supposons qu'elle est valable pour n et montrons qu'elle l'est pour

*) Leibniz Gottfried Wilhelm (1646-1716), philosophe et mathématicien allemand.

$n + 1$. Dérivons la formule (1):

$$y^{(n+1)} = u^{(n+1)}v + u^{(n)}v' + n\,[u^{(n)}v' + u^{(n-1)}v''] +$$
$$+ \frac{n(n-1)}{2!}\,[u^{(n-1)}v'' + u^{(n-2)}v'''] + \ldots$$
$$\ldots + \frac{n(n-1)\ldots(n-k+1)}{k!}\,[u^{(n-k+1)}v^{(k)} + u^{(n-k)}v^{(k+1)}] + \ldots$$
$$\ldots + u'v^{(n)} + uv^{(n+1)}.$$

En chassant les parenthèses et en réduisant les termes semblables, on obtient

$$y^{(n+1)} = u^{(n+1)}v + (n+1)\,u^{(n)}v' + \left(n + \frac{n(n-1)}{2!}\right)u^{(n-1)}v'' + \ldots$$
$$\ldots + \left[\frac{n(n-1)\ldots(n-k+2)}{(k-1)!} + \frac{n(n-1)\ldots(n-k+1)}{k!}\right]u^{(n-k+1)}v^{(k)} + \ldots$$
$$\ldots + (n+1)u'v^{(n)} + uv^{(n+1)}.$$

L'expression entre crochets peut être mise sous la forme:

$$\frac{n(n-1)\ldots(n-k+2)}{(k-1)!} + \frac{n(n-1)\ldots(n-k+1)}{k!} =$$
$$= \frac{n(n-1)(n-2)\ldots(n-k+2)(n-k+1)(n-k)(n-k-1)(n-k-2)\ldots 1}{(k-1)!\,(n-k+1)(n-k)(n-k-2)\ldots 1} +$$
$$+ \frac{n(n-1)\ldots(n-k+1)(n-k)(n-k-1)(n-k-2)\ldots 1}{k!\,(n-k)(n-k-1)(n-k-2)\ldots 1} =$$
$$= \frac{n!}{(k-1)!\,(n-k+1)!} + \frac{n!}{k!\,(n-k)!} =$$
$$= \frac{n!}{(k-1)!\,(n-k)!}\left(\frac{1}{n-k+1} + \frac{1}{k}\right) =$$
$$= \frac{n!}{(k-1)!\,(n-k)!}\,\frac{n+1}{k(n-k+1)} = \frac{(n+1)!}{k!\,(n-k+1)!} =$$
$$= \frac{(n+1)\,n(n-1)(n-2)\ldots(n-k+2)(n-k+1)\times}{\times(n-k)(n-k-1)(n-k-2)\ldots 1} =$$
$$= \frac{(n+1)\,n(n-1)(n-2)\ldots(n-k+2)}{k!}.$$

Donc

$$y^{(n+1)} = u^{(n+1)}v + (n+1)\,u^{(n)}v' + \frac{(n+1)\,n}{2!}\,u^{(n-1)}v'' + \ldots$$
$$\ldots + \frac{(n+1)\,n(n-1)\ldots(n-k+2)}{k!}\,u^{(n-k+1)}v^{(k)} + \ldots$$
$$\ldots + (n+1)\,u'v^{(n)} + uv^{(n+1)}. \blacksquare$$

Exemple 1. Calculer la dérivée d'ordre cinq de la fonction $y = x^5e^x$.

Solution. En posant $u = x^5$ et $v = e^x$, on obtient $u' = 5x^4$, $u'' = 20x^3$, $u''' = 60x^2$, $u^{(4)} = 120x$, $u^{(5)} = 120$; $v' = v'' = v''' = v^{(4)} = v^{(5)} = e^x$. En portant ces expressions dans la formule (1) pour $n = 5$ on trouve

$$y^{(5)} = 120e^x + 5 \cdot 120xe^x + \frac{5 \cdot 4}{1 \cdot 2} 60x^2e^x + \\ + \frac{5 \cdot 4 \cdot 3}{1 \cdot 2 \cdot 3} 20x^3e^x + 5 \cdot 5x^4e^x + x^5e^x = \\ = e^x (120 + 600x + 600x^2 + 200x^3 + 25x^4 + x^5).$$

Exemple 2. Calculer la dérivée d'ordre n ($n \geqslant 2$) de la fonction $y = x^2 \cos x$.

Solution. En posant $u = \cos x$ et $v = x^2$, on trouve

$$u^{(n)} = \cos \left(x + n \frac{\pi}{2}\right), \quad v' = 2x, \quad v'' = 2, \quad v''' = v^{(4)} = v^{(5)} = \ldots = 0.$$

En portant ceci dans la formule (1), on obtient

$$y^{(n)} = \cos \left(x + n \frac{\pi}{2}\right) x^2 + 2n \cos \left[x + (n-1) \frac{\pi}{2}\right] x + \\ + \frac{2n(n-1)}{1 \cdot 2} \cos \left[x + (n-2) \frac{\pi}{2}\right].$$

4. Différentielles supérieures. Considérons les différentielles supérieures. Pour la commodité on se servira aussi bien des symboles dy et dx que des symboles δy et δx pour désigner les différentielles.

Soit $f(x)$ une fonction dérivable en chaque point x d'un intervalle. Alors sa différentielle

$$dy = f'(x)\, dx,$$

que l'on appellera *différentielle du premier ordre* (ou *différentielle première*), est une fonction de deux variables: l'argument x et sa différentielle dx. Supposons que la fonction $f'(x)$ est à son tour dérivable en un point x. Traitons dx comme un facteur constant. La fonction dy est alors une fonction du seul argument x et sa différentielle en x est de la forme:

$$\delta (dy) = \delta [f'(x)\, dx] = [f'(x) \cdot dx]' \, \delta x = f''(x)\, dx\, \delta x.$$

La différentielle $\delta (dy)$ de la différentielle dy en x, prise pour $\delta x = dx$, s'appelle *différentielle d'ordre deux* (ou *différentielle seconde*) de la fonction $f(x)$ en x et se note d^2y:

$$d^2y = f''(x)\, (dx)^2.$$

La différentielle $\delta (d^2y)$ de la différentielle d^2y, prise pour $\delta x = dx$, s'appelle *différentielle d'ordre trois* (ou *différentielle troisième*) de la fonction $f(x)$ et se note d^3y, et ainsi de suite. La dif-

férentielle $\delta(d^{n-1}y)$ de la différentielle $d^{n-1}y$, prise pour $\delta x = dx$, s'appelle *différentielle d'ordre n* (ou *différentielle n-ième*) de la fonction $f(x)$ et se note $d^n y$.

Montrons que la différentielle n-ième de la fonction y est susceptible de la formule

$$d^n y = y^{(n)}(x)(dx)^n, \quad n = 1, 2, \ldots \tag{2}$$

Faisons la démonstration par récurrence. Pour $n = 1$ et $n = 2$ la formule (2) est prouvée. Supposons qu'elle est vraie pour les différentielles d'ordre $n - 1$:

$$d^{n-1}y = y^{(n-1)}(x)(dx)^{n-1},$$

et que la fonction $y^{(n-1)}(x)$ est dérivable en un point x. Alors

$$d^n y = \delta(d^{n-1}y) = \delta[y^{(n-1)}(x)(dx)^{n-1}] =$$
$$= [y^{(n-1)}(x)(dx)^{n-1}]' \delta x = y^{(n)}(x)\delta x (dx)^{n-1}.$$

En posant $\delta x = dx$, on obtient

$$d^n y = \delta(d^{n-1}y)|_{\delta x = dx} = y^{(n)}(x)(dx)^n,$$

C.Q.F.D.

La formule (2) entraîne pour tout n

$$y^{(n)}(x) = \frac{d^n y}{(dx)^n} \quad \text{ou} \quad y^{(n)}(x) = \frac{d^n y}{dx^n}.$$

c'est-à-dire que la dérivée d'ordre n de la fonction $y = f(x)$ au point x est égale au rapport de la différentielle d'ordre n de cette fonction en x à la puissance n-ième de la différentielle de l'argument.

Exemple 3. Calculer la différentielle d^3y de la fonction $y = x^4 - 3x^2 + 4$.

Solution. Une dérivation successive nous donne

$$y'(x) = 4x^3 - 6x, \quad y''(x) = 12x^2 - 6, \quad y'''(x) = 24x.$$

Donc

$$d^3y = y'''(x)(dx)^3 = 24x(dx)^3.$$

§ 11. Définition paramétrique d'une fonction et dérivation

1. Définition paramétrique d'une fonction. Soient données deux fonctions

$$x = \varphi(t), \qquad y = \psi(t) \tag{1}$$

d'une variable indépendante t, définies et continues sur le même intervalle. Si $x = \varphi(t)$ est strictement monotone, sa réciproque $t = \Phi(x)$ est univoque, continue et strictement monotone. Donc y peut être traitée comme une fonction dépendant de la variable x

par l'intermédiaire de la variable t, appelée *paramètre*:

$$y = \psi [\Phi (x)].$$

On dit alors que la fonction y de x est *définie paramétriquement* par les équations (1).

Signalons que la fonction $\psi [\Phi (x)]$ est continue en vertu du théorème de continuité d'une fonction composée.

Exemple 1. Soient $x = R \cos t$, $y = R \sin t$ $(0 \leqslant t \leqslant \pi)$. La fonction $x = R \cos t$ étant strictement décroissante sur $[0, \pi]$, ces équations définissent paramétriquement une fonction y de x. Si à partir de la première équation on exprime t en fonction de x et qu'on le porte dans la deuxième, on obtient la fonction de x cherchée sous forme explicite.

On obtient plus facilement y en fonction de x si l'on remarque que

$$x^2 + y^2 = R^2 (\cos^2 t + \sin^2 t) = R^2.$$

D'où $y = \pm\sqrt{R^2 - x^2}$. La fonction $y = R \sin t$ étant positive sur $[0, \pi]$, on retiendra le radical précédé du signe $+$, c'est-à-dire $y = \sqrt{R^2 - x^2}$.

Si $t \in [\pi, 2\pi]$, alors $y = -\sqrt{R^2 - x^2}$.

On en déduit donc que lorsque t varie sur l'intervalle $[0, 2\pi]$, les formules $x = R \cos t$ et $y = R \sin t$ définissent deux fonctions de la variable x représentées graphiquement par un cercle de rayon R.

Exemple 2. Soient $x = a \cos t$, $y = b \sin t$, $t \in [0, 2\pi]$. Ces équations sont les équations paramétriques de l'ellipse, puisque (cf. remarque du n° 1, § 7, chap. 3) l'ellipse est la transformée de son cercle principal dans une affinité orthogonale d'axe Ox et de rapport $\frac{b}{a}$. De l'exemple 1 il s'ensuit que les équations paramétriques du cercle $x^2 + y^2 = a^2$ sont les équations $x = a \cos t$, $y = a \sin t$, $t \in [0, 2\pi]$. Donc les équations paramétriques de l'ellipse se déduisent de celles du cercle en multipliant le second membre de l'équation de y par b/a, soit: $x = a \cos t$, $y = b \sin t$, $t \in [0, 2\pi]$. On peut procéder autrement. En éliminant t entre ces équations (en les résolvant par rapport à $\cos t$ et $\sin t$, en élevant au carré et en additionnant) on obtient

$$(x/a)^2 + (y/b)^2 = \cos^2 t + \sin^2 t = 1 \quad \text{ou} \quad x^2/a^2 + y^2/b^2 = 1$$

c'est-à-dire l'équation d'une ellipse.

La définition paramétrique d'une fonction revêt une signification toute particulière dans l'étude du mouvement d'un point. Si un point se déplace sur le plan, ses coordonnées x et y sont fonctions du temps t. La donnée des fonctions $x = \varphi (t)$, $y =$

$= \psi(t)$ définit complètement le mouvement du point. Sur chaque intervalle de temps où la fonction $\varphi(t)$ est strictement monotone, on peut comme précédemment définir une fonction $y = \psi[\Phi(x)]$ dont la courbe représentative est la trajectoire décrite par le point durant cet intervalle de temps. Dans le dernier exemple les fonctions décrivaient le mouvement elliptique d'un point.

2. Dérivation d'une fonction définie paramétriquement. Supposons maintenant que les fonctions $x = \varphi(t)$ et $y = \psi(t)$ sont dérivables et de plus que $\varphi'(t) \neq 0$ sur un intervalle. Cette inégalité entraîne la stricte monotonie de la fonction $x = \varphi(t)$ (cf. théorème 6.7 du chap. VI) et par suite l'univocité de la fonction réciproque $t = \Phi(x)$. D'après le théorème 5.4 de la dérivée de la fonction réciproque, la fonction $\Phi(x)$ admet la dérivée

$$\Phi'(x) = \frac{1}{\varphi'(t)}$$

et d'après le théorème 5.5 de la dérivée d'une fonction composée, la fonction $y = \psi[\Phi(x)]$ a pour dérivée

$$y'(x) = \psi'(\Phi(x))\Phi'(x).$$

Donc

$$y'(x) = \frac{\psi'(t)}{\varphi'(t)}\Big|_{t=\Phi(x)}. \qquad (2)$$

On a donc prouvé que la dérivée d'une fonction définie paramétriquement s'exprime par la formule (2).

Exemple 1. Trouver $y'(x)$ si $x = R\cos t$, $y = R\sin t$, $t \in [0, \pi]$.

Solution. La formule (2) nous donne (ici $t = \Phi(x) = \operatorname{Arc}\cos(x/R)$):

$$y'(x) = \frac{R\cos t}{-R\sin t}\Big|_{t=\operatorname{Arc}\cos(x/R)} = -\frac{\cos t}{\sqrt{1-\cos^2 t}}\Big|_{t=\operatorname{Arc}\cos(x/R)} =$$

$$= -\frac{\frac{x}{R}}{\sqrt{1-x^2/R^2}} = -\frac{x}{\sqrt{R^2-x^2}} \quad (x \neq \pm R).$$

Si l'on se sert de l'expression explicite de la fonction y, soit $y = \sqrt{R^2 - x^2}$, on obtient bien sûr le même résultat:

$$y'(x) = \frac{-2x}{2\sqrt{R^2-x^2}} = -\frac{x}{\sqrt{R^2-x^2}} \quad (x \neq \pm R).$$

Supposons que les fonctions $\varphi(t)$ et $\psi(t)$ sont deux fois dérivables en un point t. On peut alors calculer la dérivée seconde de la fonction définie paramétriquement. On remarquera que la fonction $y'(x) = \frac{\psi'(t)}{\varphi'(t)}\Big|_{t=\Phi(x)}$ est à son tour définie paramétriquement par les équations $y'_x = \frac{\psi'(t)}{\varphi'(t)} = \psi_1(t)$ et $x = \varphi(t)$.

La formule (2) nous donne donc

$$y''(x) = (y'(x))'_x = \frac{\psi'_t(t)}{\varphi'(t)}\Big|_{t=\Phi(x)} = \frac{\left(\frac{\psi'(t)}{\varphi'(t)}\right)'}{\varphi'(t)}\Big|_{t=\Phi(x)} =$$

$$= \frac{\frac{\psi''(t)\varphi'(t) - \varphi''(t)\psi'(t)}{(\varphi'(t))^2}}{\varphi'(t)}\Big|_{t=\Phi(x)} = \frac{\psi''(t)\varphi'(t) - \varphi''(t)\psi'(t)}{[\varphi'(t)]^3}\Big|_{t=\Phi(x)}.$$

On s'est servi de la règle de dérivation du quotient. Ainsi

$$y''(x) = \frac{\psi''(t)\varphi'(t) - \varphi''(t)\psi'(t)}{[\varphi'(t)]^3}\Big|_{t=\Phi(x)}. \qquad (3)$$

On calcule de façon analogue la dérivée de tout ordre de y par rapport à x.

Exemple 2. Calculer $y''(x)$ si $x = \cos t$, $y = \sin t$, $t \in [0, \pi]$.

Solution. $y'(t) = \cos t$, $y''(t) = -\sin t$; $x'(t) = -\sin t$, $x''(t) = -\cos t$. En portant dans la formule (3), on trouve

$$y''(x) = \frac{(-\sin t)(-\sin t) - (-\cos t)(\cos t)}{(-\sin t)^3}\Big|_{t=\operatorname{Arc}\cos(x/R)} =$$

$$= \frac{\sin^2 t + \cos^2 t}{(-\sin t)^3}\Big|_{t=\operatorname{Arc}\cos(x/R)} = -\frac{1}{\sin^3 t}\Big|_{t=\operatorname{Arc}\cos(x/R)} = -\frac{1}{(1-x^2/R^2)^{3/2}}.$$

CHAPITRE VI

APPLICATION DU CALCUL DIFFÉRENTIEL À L'ÉTUDE DES FONCTIONS

Dans le chapitre précédent nous avons étudié la dérivation des fonctions. Voyons maintenant les méthodes d'étude des fonctions et de construction des graphiques, méthodes qui sont largement utilisées aussi bien en théorie qu'en pratique.

§ 1. Théorèmes fondamentaux du calcul différentiel

Théorème 6.1 (Fermat *)). *Soit $f(x)$ une fonction définie sur un intervalle $]a, b[$ et présentant un extrémum en un point x_0. Si cette fonction est dérivable en x_0, sa dérivée est nulle en ce point, soit $f'(x_0) = 0$.*

Démonstration. Supposons pour fixer les idées que la fonction $f(x)$ présente un maximum en x_0, c'est-à-dire que $f(x) \leqslant f(x_0)$ pour tout $x \in]a, b[$. Ceci exprime que $\Delta y = f(x_0 + \Delta x) - f(x_0) \leqslant 0$ pour tout point $x_0 + \Delta x \in]a, b[$. Donc si $\Delta x > 0$ $(x > x_0)$, alors $\Delta y/\Delta x \leqslant 0$ et par suite

$$f'_+(x_0) = \lim_{\Delta x \to 0+} \frac{\Delta y}{\Delta x} \leqslant 0;$$

si $\Delta x < 0$, alors $\Delta y/\Delta x \geqslant 0$ et par suite

$$f'_-(x_0) = \lim_{\Delta x \to 0-} \frac{\Delta y}{\Delta x} \geqslant 0,$$

c'est-à-dire qu'en x_0 la dérivée à droite est négative et la dérivée à gauche, positive. La dérivée $f'(x_0)$ existe par hypothèse, donc $f'_+(x_0) = f'_-(x_0) = f'(x_0)$. Ce qui n'est possible que lorsque $f'_+(x_0) = f'_-(x_0) = 0$. Donc $f'(x_0) = 0$.

Le cas où $f(x)$ présente un minimum en x_0 se traite de façon analogue. ■

Le théorème de Fermat admet l'interprétation géométrique suivante: si une fonction $f(x)$ dérivable en un point x_0 présente un extrémum en ce point, la tangente à sa courbe représentative au point $(x_0; f(x_0))$ est parallèle à l'axe Ox (fig. 69).

---

*) Fermat Pierre (1601-1665), mathématicien français.

Remarque. Le théorème est mis en défaut si la fonction $f(x)$ est envisagée sur un intervalle fermé $[a, b]$. Par exemple, sur $[0, 1]$ la fonction $f(x) = x$ présente un minimum en $x = 0$ et un maximum en $x = 1$, mais sa dérivée n'est nulle en aucun de ces points (fig. 70).

Théorème 6.2 (Rolle *)). *Soit $f(x)$ une fonction:*

1° *définie et continue sur $[a, b]$;*

2° *dérivable sur $]a, b[$;*

3° $f(a) = f(b)$.

Il existe alors un point $c \in]a, b[$ en lequel $f'(c) = 0$.

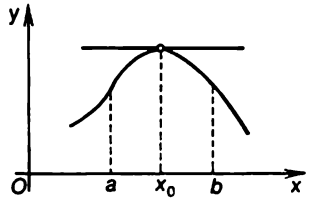


Fig. 69

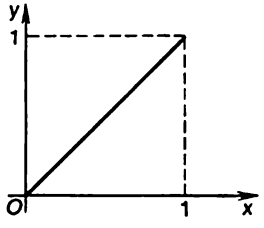


Fig. 70

Démonstration. La fonction $f(x)$ étant continue sur $[a, b]$, elle prend d'après le deuxième théorème de Weierstrass une valeur maximale M et une valeur minimale m, c'est-à-dire qu'il existe deux points $x_1, x_2 \in [a, b]$ tels que $f(x_1) = m$, $f(x_2) = M$ et

$$m \leqslant f(x) \leqslant M.$$

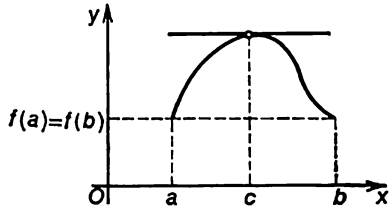


Fig. 71

On distinguera deux cas: 1) $M = m$; 2) $m < M$.

Dans le premier, $f(x) = \text{const} = M = m$. Donc la dérivée $f'(x)$ s'annule en tout point de $[a, b]$ et le théorème est prouvé.

Dans le second, puisque $f(a) = f(b)$, l'une au moins des valeurs m et M n'est pas prise aux bornes de $[a, b]$, c'est-à-dire qu'il existe un point $c \in]a, b[$ en lequel la fonction $f(x)$ prend sa valeur maximale ou minimale. Dans ce cas la fonction $f(x)$ est dérivable en c et le théorème de Fermat nous dit que $f'(c) = 0$. ■

La signification géométrique du théorème de Rolle est la suivante: le graphique d'une fonction $f(x)$ continue sur $[a, b]$, dérivable sur $]a, b[$ et telle que $f(a) = f(b)$, présente un point $(c; f(c))$ en lequel la tangente est parallèle à l'axe Ox (fig. 71). Sur la figure 71 la fonction $f(x)$ présente un maximum en c.

*) Rolle Michel (1652-1719), mathématicien français.

Signalons que les trois conditions du théorème de Rolle sont toutes essentielles. Pour s'en assurer il suffit de citer des exemples de fonctions pour lesquelles sont remplies deux conditions seulement du théorème et dont les dérivées ne s'annulent en aucun point. Ainsi par exemple la fonction $f(x) = x$, $x \in [0, 1]$ (cf.

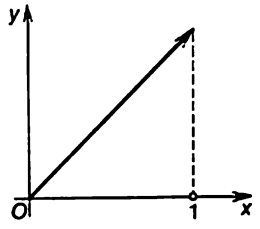


Fig. 72

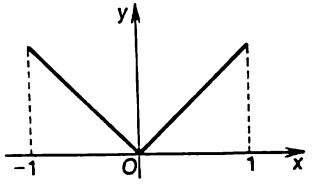


Fig. 73

fig. 70) satisfait les conditions 1° et 2° mais pas la condition 3° et il n'existe pas de point c en lequel $f'(x) = 0$. Considérons encore deux exemples.

La fonction

$$f(x) = \begin{cases} x & \text{si} \quad 0 \leqslant x < 1, \\ 0 & \text{si} \quad x = 1 \end{cases}$$

dont le graphique est représenté sur la figure 72 vérifie les conditions 2° et 3°, mais pas la condition 1°. La fonction $f(x) = |x|$, $x \in [-1, 1]$ (fig. 73) vérifie les conditions 1° et 3°, mais pas la condition 2°. Pour ces fonctions il n'existe pas non plus de point c en lequel leur dérivée fût nulle.

Signalons qu'en mathématiques on vérifie qu'une condition est essentielle en construisant des exemples dans lesquels la non-réalisation de cette condition met en défaut le théorème.

Théorème 6.3 (Lagrange *)). *Si $f(x)$ est une fonction:*
1° *définie et continue sur $[a, b]$;*
2° *dérivable sur $]a, b[$, il existe un point $c \in]a, b[$ tel que l'on a*

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

Démonstration. Considérons la fonction auxiliaire

$$F(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a),$$

sur $[a, b]$. Cette fonction remplit les trois conditions du théorème de Rolle:

*) Lagrange Joseph Louis (1736-1813), mathématicien français.

1) $F(x)$ est définie et continue sur $[a, b]$ (en tant que différence de deux fonctions continues: la fonction $f(x)$ et la fonction $f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$);

2) $F(x)$ est dérivable sur $]a, b[$, c'est-à-dire qu'à l'intérieur de $[a, b]$ elle admet une dérivée égale à $F'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$;

3) $F(a) = 0$, $F(b) = 0$ d'où $F(a) = F(b)$.

D'après le théorème de Rolle il existe donc un point $c \in]a, b[$ tel que $F'(c) = 0$, c'est-à-dire que $f'(c) - \frac{f(b)-f(a)}{b-a} = 0$. D'où $f'(c) = \frac{f(b)-f(a)}{b-a}$. ■

Etablissons la signification géométrique du théorème de Lagrange (fig. 74). La quantité $\frac{f(b)-f(a)}{b-a}$ est le coefficient directeur de la sécante passant par les points $M_1(a; f(a))$ et $M_2(b; f(b))$ de la courbe $y = f(x)$, et $f'(c)$, le coefficient directeur de la tangente à la courbe $y = f(x)$ au point $(c; f(c))$. Le théorème de Lagrange affirme qu'il existe au moins un point c tel que la tangente au point $(c; f(c))$ est parallèle à la sécante M_1M_2.

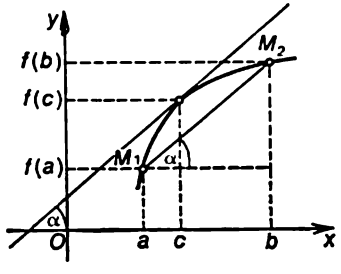


Fig. 74

Remarque 1. L'égalité

$$f(b) - f(a) = f'(c)(b-a), \quad a < c < b, \qquad (1)$$

s'appelle *formule de Lagrange* ou *formule des accroissements finis.*

Remarque 2. Le point c étant compris entre a et b, on a $c = a + \theta(b-a)$, où $0 < \theta < 1$. Ceci nous permet d'écrire la formule de Lagrange sous la forme suivante:

$$f(b) - f(a) = f'(a + \theta(b-a))(b-a), \text{ où } 0 < \theta < 1.$$

Remarque 3. Si l'on pose $a = x$, $b = x + \Delta x$, on obtient

$$f(x + \Delta x) - f(x) = f'(x + \theta\Delta x)\,\Delta x, \text{ où } 0 < \theta < 1.$$

Cette représentation de la formule de Lagrange est souvent plus commode que la forme (1).

On verra dans la suite que le théorème de Lagrange intervient dans la démonstration de nombreux théorèmes et formules d'analyse.

Théorème 6.4 (Cauchy). *Soient $f(x)$ et $g(x)$ des fonctions continues sur $[a, b]$ et dérivables sur $]a, b[$. Supposons que $g'(x) \neq 0$.*

Il existe alors un point $c \in]a, b[$ *tel que l'on a*

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(c)}{g'(c)}. \tag{2}$$

Démonstration. Montrons tout d'abord que $g(b) \neq g(a)$, c'est-à-dire que la formule (2) a un sens. En effet, si l'on admet que $g(b) = g(a)$, le théorème de Rolle nous dit qu'il existe un point $\xi \in]a, b[$ en lequel $g'(\xi) = 0$. Or ceci contredit la condition $g'(x) \neq 0$ sur $]a, b[$. Passons à la démonstration de la formule (2).

Considérons la fonction auxiliaire

$$F(x)=f(x)-f(a)-\frac{f(b)-f(a)}{g(b)-g(a)}[g(x)-g(a)]$$

sur $[a, b]$. Il est immédiat de remarquer que la fonction $F(x)$ vérifie les conditions du théorème de Rolle sur $[a, b]$. En effet, $F(x)$ est continue sur $[a, b]$, dérivable sur $]a, b[$ et de plus $F(a) = F(b) = 0$. D'après le théorème de Rolle, il existe un point $c \in]a, b[$ tel que $F'(c) = 0$.

Comme $F'(x)=f'(x)-\frac{f(b)-f(a)}{g(b)-g(a)}g'(x)$, on a

$$F'(c)=f'(c)-\frac{f(b)-f(a)}{g(b)-g(a)}g'(c)=0.$$

D'où la formule (2), puisque $g'(c) \neq 0$. ■

Remarque. La formule (2) s'appelle *formule de Cauchy* ou *formule généralisée des accroissements finis.*

§ 2. Levée des indéterminations. Règle de l'Hôpital

1. Levée d'une indétermination de la forme $\frac{0}{0}$. On dira que le rapport $\frac{f(x)}{g(x)}$ est une indétermination de la forme $\frac{0}{0}$ lorsque $x \to a$ si

$$\lim_{x\to a} f(x)=\lim_{x\to a} g(x)=0.$$

Lever cette indétermination c'est calculer la limite $\lim\limits_{x\to a}\frac{f(x)}{g(x)}$ si elle existe ou montrer qu'elle n'existe pas. Le théorème suivant nous fournit une règle pour lever les indéterminations de la forme $\frac{0}{0}$.

Théorème 6.5 (l'Hôpital *)). *Soient $f(x)$ et $g(x)$ des fonctions définies et dérivables au voisinage d'un point a, sauf éventuellement en ce point a. Supposons que* $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$ **) *et* $g'(x) \neq 0$ *dans ce voisinage du point a. Si le rapport* $\frac{f'(x)}{g'(x)}$ *tend vers une limite (finie ou infinie) lorsque* $x \to a$, *il en est de même du rapport* $\frac{f(x)}{g(x)}$ *et de plus*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

Démonstration. Soit $\{x_n\}$ une suite quelconque de valeurs de x tendant vers $a \neq x_n$. Prolongeons les fonctions $f(x)$

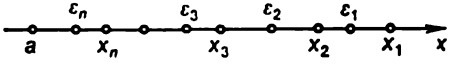


Fig. 75

et $g(x)$ par continuité au point a en posant $f(a) = g(a) = 0$. Il est alors évident que $f(x)$ et $g(x)$ sont continues sur $[a, x_n]$, dérivables sur $]a, x_n[$ et par hypothèse $g'(x) \neq 0$. Donc les fonctions $f(x)$ et $g(x)$ remplissent toutes les conditions du théorème de Cauchy sur $[a, x_n]$, autrement dit $[a, x_n]$ contient un point ξ_n tel que

$$\frac{f(x_n) - f(a)}{g(x_n) - g(a)} = \frac{f'(\xi_n)}{g'(\xi_n)}, \quad \xi_n \in]a, x_n[.$$

Mais $f(a) = g(a) = 0$, donc

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(\xi_n)}{g'(\xi_n)}, \quad \xi_n \in]a, x_n[. \tag{1}$$

Faisons tendre n vers ∞ dans la formule (1). Il est alors évident que $\xi_n \to a$ lorsque $n \to \infty$ (fig. 75). Puisque $\frac{f'(x)}{g'(x)}$ possède une limite lorsque x tend vers a, le second membre de la formule (1) possède, lorsque $n \to \infty$, une limite qui est égale à $\lim_{x\to a} \frac{f'(x)}{g'(x)}$. Donc, lorsque $n \to \infty$, le premier membre de la formule (1) admet aussi une limite et de plus

$$\lim_{n\to\infty} \frac{f(x_n)}{g(x_n)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

*) L'Hôpital Guillaume François (1661-1704), mathématicien français.
**) Le théorème est vrai dans le cas aussi où $x \to a_-$ ou $x \to a_+$.

Puisque $\{x_n\}$ est une suite arbitraire de valeurs de x tendant vers a, on en déduit que $\frac{f(x)}{g(x)}$ possède une limite lorsque $x \to a$ et

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}. \blacksquare$$

Ce théorème s'appelle généralement *règle de l'Hôpital.*

Remarque 1. Si les dérivées $f'(x)$ et $g'(x)$ remplissent les mêmes conditions que les fonctions $f(x)$ et $g(x)$, on peut appliquer la règle de l'Hôpital à deux reprises. On obtient alors

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)} = \lim_{x \to a} \frac{f''(x)}{g''(x)}.$$

Remarque 2. Le théorème reste en vigueur dans le cas où $x \to \infty$, $x \to +\infty$ et $x \to -\infty$. En effet, supposons par exemple que $\lim_{x \to \infty} f(x) = \lim_{x \to \infty} g(x) = 0$ et que $\frac{f'(x)}{g'(x)}$ possède une limite (finie ou infinie) lorsque $x \to \infty$. Faisons la substitution $x = 1/t$; alors $t \to 0$ lorsque $x \to \infty$ et $f(x) = f(1/t) \to 0$, $g(x) = g(1/t) \to 0$. En appliquant le théorème 6.5 et la règle de dérivation d'une fonction composée aux fonctions $f(1/t)$ et $g(1/t)$, on obtient

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = \lim_{t \to 0} \frac{f(1/t)}{g(1/t)} = \lim_{t \to 0} \frac{f'(1/t)(-1/t^2)}{g'(1/t)(-1/t^2)} = \lim_{t \to 0} \frac{f'(1/t)}{g'(1/t)} = \lim_{x \to \infty} \frac{f'(x)}{g'(x)}.$$

Considérons quelques exemples.

1. $\lim_{x \to 0} \frac{1 - \cos x}{x^2} = \lim_{x \to 0} \frac{\sin x}{2x} = \frac{1}{2} \lim_{x \to 0} \frac{\sin x}{x} = \frac{1}{2} \cdot 1 = \frac{1}{2}$.

2. $\lim_{x \to 0} \frac{x - \sin x}{x^3} = \lim_{x \to 0} \frac{1 - \cos x}{3x^2} = \lim_{x \to 0} \frac{\sin x}{6x} =$

$= \frac{1}{6} \lim_{x \to 0} \frac{\sin x}{x} = \frac{1}{6} \cdot 1 = \frac{1}{6}$.

3. $\lim_{x \to 0} \frac{e^x - 1}{x} = \lim_{x \to 0} \frac{e^x}{1} = \lim_{x \to 0} e^x = 1$.

2. **Levée d'une indétermination de la forme $\frac{\infty}{\infty}$.** On dira que le rapport $\frac{f(x)}{g(x)}$ est une indétermination de la forme $\frac{\infty}{\infty}$ lorsque $x \to a$ si

$$\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = \infty, \quad +\infty \quad \text{ou} \quad -\infty.$$

Cette indétermination est justiciable d'une proposition identique au théorème 6.5, plus exactement: ce théorème reste va-

lable si l'on y remplace la condition $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$ par la condition $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = \infty$.

Considérons quelques exemples.

1. $\lim\limits_{x\to\infty} \dfrac{\ln x}{x} = \lim\limits_{x\to+\infty} \dfrac{(1/x)}{1} = \lim\limits_{x\to+\infty} \dfrac{1}{x} = 0.$

2. $\lim\limits_{x\to+\infty} \dfrac{x^n}{e^x} = \lim\limits_{x\to+\infty} \dfrac{nx^{n-1}}{e^x} = \lim\limits_{x\to+\infty} \dfrac{n(n-1)x^{n-2}}{e^x} = \ldots$

$$\ldots = \lim_{x\to\infty} \frac{n!}{e^x} = 0.$$

3. $\lim\limits_{x\to 0+} \sqrt{x} \ln x = \lim\limits_{x\to 0+} \dfrac{\ln x}{x^{-1/2}} =$

$$= \lim_{x\to 0+} \frac{1/x}{(-1/2)x^{-3/2}} = -2 \lim_{x\to 0+} \sqrt{x} = 0.$$

3. Levée des autres formes d'indétermination. Les indéterminations de la forme $0\cdot\infty$ et $\infty - \infty$ peuvent être ramenées aux indéterminations de la forme $\frac{0}{0}$ et $\frac{\infty}{\infty}$. Montrons-le sur des exemples.

Exemple 1. Trouver $\lim\limits_{x\to 0+} x \ln x$.

Solution. On a une indétermination de la forme $0\cdot\infty$. Mais $x \ln x = \dfrac{\ln x}{1/x}$, c'est-à-dire qu'on a affaire à une indétermination $\frac{\infty}{\infty}$. En appliquant la règle de l'Hôpital, on trouve

$$\lim_{x\to 0+} x \ln x = \lim_{x\to 0+} \frac{(\ln x)'}{(1/x)'} = \lim_{x\to 0+} \frac{1/x}{-1/x^2} = -\lim_{x\to 0+} x = 0.$$

Exemple 2. Trouver $\lim\limits_{x\to\pi/2} (\sec x - \operatorname{tg} x)$.

Solution. On a une indétermination de la forme $\infty - \infty$. Mais $\sec x - \operatorname{tg} x = \dfrac{1}{\cos x} - \dfrac{\sin x}{\cos x} = \dfrac{1-\sin x}{\cos x}$. On s'est donc ramené à une indétermination de la forme $\frac{0}{0}$.

La règle de l'Hôpital nous donne

$$\lim_{x\to\pi/2} (\sec x - \operatorname{tg} x) = \lim_{x\to\pi/2} \left(\frac{1-\sin x}{\cos x}\right) = \lim_{x\to\pi/2} \frac{-\cos x}{-\sin x} = 0.$$

Considérons enfin les indéterminations de la forme 0^0, 1^∞ et ∞^0. On rencontre ces indéterminations lorsqu'on étudie des fonctions de la forme $y = f(x)^{g(x)}$ dans lesquelles $f(x)$ tend respectivement vers 0, 1 et ∞, et $g(x)$, respectivement vers 0, ∞ et 0

lorsque $x \to a$. L'identité

$$f(x)^{g(x)} = e^{g(x) \ln f(x)}$$

nous ramène à l'indétermination $0 \cdot \infty$ qui a déjà été étudiée.

Exemple 3. Calculer $\lim\limits_{x \to 0+} x^x$.

Solution. On a une indétermination de la forme 0^0. Mais $x^x = e^{x \ln x}$. A l'exposant on a une indétermination de la forme $0 \cdot \infty$ qui a déjà été étudiée (cf. exemple 1). Donc

$$\lim_{x \to 0+} x^x = \lim_{x \to 0+} e^{x \ln x} = e^{\lim\limits_{x \to 0+} x \ln x} = e^0 = 1.$$

Exemple 4. Calculer $\lim\limits_{x \to 0} (1+x^2)^{\frac{1}{e^x - 1 - x}}$.

Solution. On a une indétermination de la forme 1^∞. Mais

$$(1+x^2)^{1/(e^x - 1 - x)} = e^{[\ln (1+x^2)]/(e^x - 1 - x)}.$$

A l'exposant on a une indétermination de la forme $\frac{0}{0}$. La règle de l'Hôpital nous donne

$$\lim_{x \to 0} \frac{\ln (1+x^2)}{e^x - 1 - x} = \lim_{x \to 0} \frac{2x/(1+x^2)}{e^x - 1} =$$

$$= \lim_{x \to 0} \frac{2x}{(e^x - 1)(1+x^2)} = \lim_{x \to 0} \frac{2}{e^x (1+x^2) + (e^x - 1) 2x} = \frac{2}{1} = 2.$$

Donc

$$\lim_{x \to 0} (1+x^2)^{\frac{1}{e^x - 1 - x}} = e^{\lim\limits_{x \to 0} \frac{\ln (1+x^2)}{e^x - 1 - x}} = e^2.$$

Exemple 5. Trouver $\lim\limits_{x \to \pi/2} (\operatorname{tg} x)^{2 \cos x}$.

Solution. On a une indétermination de la forme ∞^0. Mais

$$(\operatorname{tg} x)^{2 \cos x} = e^{2 \cos x \ln \operatorname{tg} x} = e^{\frac{2 \ln \operatorname{tg} x}{1/(\cos x)}}.$$

A l'exposant on a une indétermination de la forme $\frac{\infty}{\infty}$. La règle de l'Hôpital nous donne

$$\lim_{x \to \pi/2} \frac{2 \ln \operatorname{tg} x}{1/(\cos x)} = 2 \lim_{x \to \pi/2} \frac{\ln \operatorname{tg} x}{\sec x} = 2 \lim_{x \to \pi/2} \frac{1/(\operatorname{tg} x) \sec^2 x}{\sec x \operatorname{tg} x} =$$

$$= 2 \lim_{x \to \pi/2} \frac{\sec x}{\operatorname{tg}^2 x} = 2 \lim_{x \to \pi/2} \frac{\sec x \operatorname{tg} x}{2 \operatorname{tg} x \sec^2 x} = \lim_{x \to \pi/2} \cos x = 0.$$

Donc

$$\lim_{x \to \pi/2} (\operatorname{tg} x)^{2 \cos x} = e^{\lim\limits_{x \to \pi/2} 2 \cos x \ln \operatorname{tg} x} = e^0 = 1.$$

§ 3. Formule de Taylor

Etudions maintenant l'une des plus importantes formules d'analyse mathématique dont les applications sont innombrables tant en analyse que dans les disciplines connexes.

1. Formule de Taylor

Théorème 6.6 (Taylor *)). *Soit $f(x)$ une fonction $n+1$ fois dérivable* **) *dans un voisinage d'un point a et dans a, et soit x une valeur quelconque de ce voisinage ($x \neq a$). Entre les points a et x il existe alors un point ξ tel que l'on a la formule*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots$$

$$\ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}. \qquad (1)$$

Démonstration. Désignons par $\varphi(x, a)$ le polynôme de degré n en x du second membre de la formule (1), soit

$$\varphi(x, a) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

(On l'appelle polynôme de Taylor de degré n de la fonction $f(x)$.)

Désignons par $R_{n+1}(x)$ la différence

$$R_{n+1}(x) = f(x) - \varphi(x, a).$$

On prouvera le théorème lorsqu'on aura établi que

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}, \quad a < \xi < x.$$

Fixons une valeur quelconque x dans le voisinage envisagé. Admettons pour fixer les idées que $x > a$. Définissons une variable t sur $[a, x]$ et considérons la fonction auxiliaire

$$F(t) = f(x) - \varphi(x, t) - \frac{(x-t)^{n+1} R_{n+1}(x)}{(x-a)^{n+1}} \qquad (2)$$

sur $[a, x]$.

La fonction $F(t)$ remplit sur $[a, x]$ toutes les conditions du théorème de Rolle : 1) de la formule (2) et des conditions imposées à la fonction $f(x)$ il s'ensuit que $F(t)$ est continue et dérivable sur $[a, x]$, puisque $f(t)$ et toutes ses dérivées jusqu'à l'ordre n le sont sur $[a, x]$;

2) en posant $t = a$ dans (2), on obtient

$$F(a) = f(x) - \varphi(x, a) - R_{n+1}(x) =$$

$$= R_{n+1}(x) - R_{n+1}(x) = 0.$$

*) Taylor Brook (1685-1731), mathématicien anglais.

**) Il s'ensuit que la fonction $f(x)$ et toutes ses dérivées jusqu'à l'ordre n sont continues et dérivables dans ce voisinage.

En posant $t = x$ dans (2), on a

$$F(x) = f(x) - f(x) - \frac{f'(x)}{1!}(x-x) - \frac{f''(x)}{2!}(x-x)^2 - \ldots$$
$$\ldots - \frac{f^{(n)}(x)}{n!}(x-x)^n - \frac{(x-x)^{n+1}R_{n+1}(x)}{(x-a)^{n+1}} = 0.$$

Donc la condition $F(a) = F(x)$ est remplie.

D'après le théorème de Rolle, l'intervalle $[a, x]$ contient un point ξ tel que

$$F'(\xi) = 0. \tag{3}$$

Calculons la dérivée $F'(t)$. Une dérivation de la formule (2) par rapport à t nous donne

$$F'(t) = -f'(t) + \frac{f'(t)}{1!} - \frac{f''(t)}{1!}(x-t) + \frac{f''(t)}{2!}2(x-t) -$$
$$- \frac{f'''(t)}{2!}(x-t)^2 + \ldots + \frac{f^{(n)}(t)}{n!}n(x-t)^{n-1} -$$
$$- \frac{f^{(n+1)}(t)}{n!}(x-t)^n + \frac{(n+1)(x-t)^n R_{n+1}(x)}{(x-a)^{n+1}}.$$

Il est immédiat de remarquer que tous les termes du second membre à l'exception des deux derniers s'annulent mutuellement. Donc

$$F'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n + \frac{(n+1)(x-t)^n R_{n+1}(x)}{(x-a)^{n+1}}. \tag{4}$$

En faisant $t = \xi$ dans (4) et en se servant de la relation (3), on obtient

$$F'(\xi) = -\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n + \frac{(n+1)(x-\xi)^n R_{n+1}(x)}{(x-a)^{n+1}} = 0,$$

d'où

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}. \quad \blacksquare$$

La formule (1) s'appelle *formule de Taylor* et l'expression de $R_{n+1}(x)$, *reste de Lagrange*. Ce reste peut être mis sous une autre forme. Puisque $\xi \in]a, x[$, il existe un nombre θ compris entre 0 et 1 tel que $\xi = a + \theta(x - a)$ et le reste prend la forme

$$R_{n+1}(x) = \frac{f^{(n+1)}[a+\theta(x-a)]}{(n+1)!}(x-a)^{n+1}, \quad 0 < \theta < 1.$$

Cette forme du reste est le plus fréquemment utilisée dans les applications.

2. **Autre écriture de la formule de Taylor et du reste.** La formule de Taylor (1) est souvent représentée sous une autre forme. En posant $a = x_0$, $x - a = \Delta x$ et $x = x_0 + \Delta x$ dans la formule (1),

on obtient

$$f(x_0+\Delta x)-f(x_0)=\frac{f'(x_0)}{1!}\Delta x+\frac{f''(x_0)}{2!}(\Delta x)^2+\ldots$$

$$\ldots+\frac{f^n(x_0)}{n!}(\Delta x)^n+\frac{f^{(n+1)}(x_0+\theta\Delta x)}{(n+1)!}(\Delta x)^{n+1},\quad 0<\theta<1. \qquad (5)$$

Pour $n=0$ on déduit de (5) la formule de Lagrange

$$f(x_0+\Delta x)-f(x_0)=f'(x_0+\theta\Delta x)\,\Delta x.$$

Montrons que si la fonction $f^{(n+1)}(x)$ est bornée au voisinage d'un point a, le reste $R_{n+1}(x)$ est un infiniment petit d'ordre supérieur à $(x-a)^n$ lorsque $x\to a$:

$$\lim_{x\to a}\frac{R_{n+1}(x)}{(x-a)^n}=\lim_{x\to a}\frac{f^{(n+1)}(\xi)}{(n+1)!}\,\frac{(x-a)^{n+1}}{(x-a)^n}=\lim_{x\to a}\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)=0,$$

puisque la fonction $f^{(n+1)}(\xi)$ est bornée et $(x-a)\to 0$ lorsque $x\to a$. Donc

$$R_{n+1}(x)=o\,[(x-a)^n] \text{ lorsque } x\to a. \qquad (6)$$

La formule (6) s'appelle *reste de Peano* *).

3. Formule de Maclaurin. On appelle *formule de Maclaurin* **) la formule de Taylor (1) pour $a=0$:

$$f(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+\frac{f^{(n)}(0)}{n!}x^n+R_{n+1}(x).$$

Le reste est

1) $R_{n+1}(x)=\dfrac{f^{(n+1)}(\theta x)}{(n+1)!}x^{n+1}$, $0<\theta<1$, sous la forme de Lagrange;

2) $R_{n+1}(x)=o(x^n)$ sous la forme de Peano.

4. Développement de certaines fonctions élémentaires à l'aide de la formule de Maclaurin.

1) $f(x)=e^x$. Puisque

$$f(x)=f'(x)=f''(x)=\ldots=f^{(n+1)}(x)=e^x,$$
$$f(0)=f'(0)=f''(0)=\ldots=f^{(n+1)}(0)=1,$$

la formule de Maclaurin est

$$e^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots+\frac{x^n}{n!}+o(x^n). \qquad (7)$$

*) Peano Giuseppe (1858-1932), mathématicien italien.
**) Maclaurin Colin (1698-1746), mathématicien écossais.

2) $f(x)=\sin x$. Comme

$$f^{(n)}(x)=\sin\left(x+n\frac{\pi}{2}\right),\quad f^{(n)}(0)=\sin\left(n\frac{\pi}{2}\right)=$$

$$=\begin{cases}0 & \text{pour } n \text{ pair},\\ (-1)^{(n-1)/2} & \text{pour } n \text{ impair},\end{cases}$$

la formule de Maclaurin est

$$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ldots+(-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!}+o(x^{2n}). \quad (8)$$

3) $f(x)=\cos x$. Etant donné que

$$f^{(n)}(x)=\cos\left(x+n\frac{\pi}{2}\right),$$

$$f^{(n)}(0)=\cos\left(n\frac{\pi}{2}\right)=\begin{cases}0 & \text{pour } n \text{ pair},\\ (-1)^{n/2} & \text{pour } n \text{ impair},\end{cases}$$

la formule de Maclaurin est

$$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\ldots+(-1)^n\frac{x^{2n}}{(2n)!}+o(x^{2n+1}). \quad (9)$$

Dans la formule (8) le reste est écrit sous la forme $o(x^{2n})$ et non pas sous la forme $o(x^{2n-1})$, puisque le terme suivant est nul (ceci concerne également la formule (9)).

4) $f(x)=(1+x)^\alpha$, où α est un nombre réel. Puisque

$$f^{(n)}(x)=\alpha(\alpha-1)\ldots(\alpha-n+1)(1+x)^{\alpha-n},$$

$$f^{(n)}(0)=\alpha(\alpha-1)\ldots(\alpha-n+1),$$

la formule de Maclaurin s'écrit

$$(1+x)^\alpha=1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+\ldots+\frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}x^n+$$

$$+R_{n+1}(x),$$

où le reste de Lagrange vaut

$$R_{n+1}(x)=\frac{\alpha(\alpha-1)\ldots(\alpha-n)}{(n+1)!}(1+\theta x)^{\alpha-(n+1)}x^{n+1},\quad 0<\theta<1.$$

Dans le cas particulier où $\alpha=n$ est un entier naturel, $f^{(n+1)}(x)=0$, donc $R_{n+1}(x)=0$ et l'on obtient la classique *formule du binôme de Newton*

$$(1+x)^n=1+\frac{n}{1!}x+\frac{n(n-1)}{2!}x^2+\ldots+x^n. \quad (10)$$

Les développements ci-dessus montrent que la formule de Maclaurin permet d'approcher les fonctions avec une certaine précision par des polynômes. Les polynômes se prêtent bien aux opérations arithmétiques, on peut calculer leurs valeurs en tout point, etc. Les formules de Taylor et de Maclaurin permettent d'approcher

des fonctions plus complexes par des polynômes. Ces formules possèdent par ailleurs un vaste champ d'applications. On se bornera à deux d'entre elles.

5. Application de la formule de Maclaurin au calcul des dérivées. La formule de Taylor est un outil très efficace de calcul des limites de fonctions, limites auxquelles on est souvent confronté en étudiant les fonctions.

Exemple 1. Trouver $\lim\limits_{x\to 0}\dfrac{\sin x-x}{x^3}$.

Solution. Pour $n=2$ la formule (8) nous donne

$$\lim_{x\to 0}\frac{x-\frac{x^3}{3!}+o(x^4)-x}{x^3}=\lim_{x\to 0}\frac{-\frac{1}{3!}+\frac{o(x^4)}{x^3}}{1}=$$

$$=-\frac{1}{3!}+\lim_{x\to 0}\frac{o(x^4)}{x^3}=-\frac{1}{3!}+0=-\frac{1}{6}.$$

Exemple 2. Trouver $\lim\limits_{x\to 0}\dfrac{e^{-x^2/2}-\cos x}{x^3\sin x}$.

Solution. Les formules (7), (8) et (9) entraînent

$$\lim_{x\to 0}\frac{e^{-x^2/2}-\cos x}{x^3\sin x}=\lim_{x\to 0}\frac{1-x^2/2+x^4/8+o(x^4)-1+x^2/2-x^4/24}{x^3(x+o(x))}=$$

$$=\lim_{x\to 0}\frac{x^4/8-x^4/24+o(x^4)}{x^4+o(x^4)}=\lim_{x\to 0}\frac{1/8-1/24+\alpha(x)}{1+\alpha(x)}=\frac{1}{8}-\frac{1}{24}=\frac{1}{12}$$

(par $\alpha(x)$ on a désigné l'infiniment petit $o(x^4)/x^4$, $x\to 0$).

6. Calcul du nombre e. Au n° 2, § 3, chap. II, on a défini le nombre e comme la limite de la suite $\{(1+1/n)^n\}$ et on a obtenu l'encadrement grossier $2\leqslant e\leqslant 3$.

Montrons comment on peut calculer le nombre e avec la précision voulue. A cet effet, écrivons la formule (7) avec un reste de Lagrange:

$$e^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}+\frac{e^{\theta x}}{(n+1)!}x^{n+1},\quad 0<\theta<1. \tag{11}$$

Si l'on remplace la fonction e^x par son polynôme de Taylor de degré n, on obtient l'égalité approchée

$$e^x\approx 1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}, \tag{12}$$

dont l'erreur absolue est

$$|R_{n+1}(x)|=\frac{e^{\theta x}}{(n+1)!}|x|^{n+1},\quad 0<\theta<1.$$

Si l'on considère la fonction e^x sur $[-1, 1]$, on obtient

$$|R_{n+1}(x)| \leqslant \frac{e^{\theta x}}{(n+1)!} < \frac{3}{(n+1)!}. \qquad (13)$$

En posant $x = 1$ dans (12), on obtient la valeur approchée de e:

$$e \approx 1 + 1 + \frac{1}{2!} + \ldots + \frac{1}{n!}.$$

L'erreur absolue est inférieure à $3/(n+1)!$ Si l'on veut calculer e à 0,001 près, il faudra définir n à partir de l'inégalité $\frac{3}{(n+1)!} < 0{,}001$ ou $(n+1)! > 3000$. D'où $n = 6$. Donc

$$e \approx 2 + \frac{1}{2!} + \ldots + \frac{1}{6!} = 2{,}718$$

à 0,001 près.

Si l'on se sert de la formule de Maclaurin, on peut donc calculer e avec n'importe quelle précision avec un algorithme basé sur les formules (11) et (13) et facilement réalisable sur ordinateur.

§ 4. Etude des fonctions et construction de leurs courbes représentatives

1. Critère de monotonie d'une fonction.

Théorème 6.7. *Soit $f(x)$ une fonction dérivable sur $]a, b[$. Si $f'(x) \geqslant 0$ (resp. $\leqslant 0$) sur $]a, b[$, la fonction $f(x)$ est croissante* (resp. *décroissante*) *sur $]a, b[$.*

Démonstration. Pour fixer les idées on traitera le cas où $f'(x) \geqslant 0$. Soient x_1 et x_2 deux points quelconques de $]a, b[$ tels que $x_1 < x_2$; sur le segment $[x_1, x_2]$ sont alors remplies toutes les conditions du théorème de Lagrange en vertu duquel

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1), \quad c \in]x_1, x_2[.$$

Par hypothèse, $f'(c) \geqslant 0$, $x_2 - x_1 > 0$, donc $f(x_2) - f(x_1) \geqslant 0$ ou $f(x_2) \geqslant f(x_1)$, c'est-à-dire que la fonction $f(x)$ est croissante sur $]a, b[$.

Le cas $f'(x) \leqslant 0$ se traite de façon analogue. ■

Remarque. On démontre de même que si $f'(x) > 0$ (resp. $f'(x) < 0$) sur $]a, b[$, la fonction $f(x)$ est strictement croissante (resp. strictement décroissante) sur $]a, b[$.

2. Extrémum local d'une fonction.

Définition. On dit qu'une fonction $f(x)$ présente un *maximum* (resp. un *minimum*) *local strict* en un point x_0 si pour tous les points x d'un δ-voisinage de x_0 on a $f(x) < f(x_0)$ (resp. $f(x) > f(x_0)$) pour $x \neq x_0$ (fig. 76).

Maximum local et minimum local sont regroupés sous le nom générique d'*extrémum local*.

De la définition il ressort que la notion d'extrémum revêt un caractère local en ce sens que l'inégalité $f(x) < f(x_0)$ (resp. $f(x) > f(x_0)$) peut ne pas être réalisée pour toutes les valeurs x du domaine de définition de $f(x)$ mais seulement dans un voisinage du point x_0. Il est évident qu'une fonction peut présenter plu-

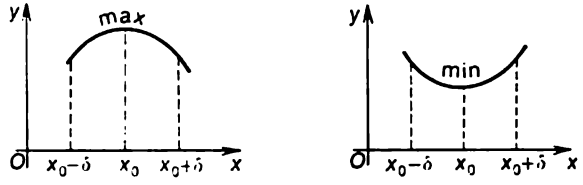


Fig. 76

sieurs maximums locaux et plusieurs minimums locaux et il est même possible que tel maximum local soit inférieur à un minimum local.

Théorème 6.8 (condition nécessaire d'extrémum local). *Si une fonction $f(x)$ est différentiable et présente un extrémum local en un point x_0, alors $f'(x_0) = 0$.*

Démonstration. Vu que la fonction $f(x)$ présente un extrémum local en x_0, il existe un intervalle $]x_0 - \delta, x_0 + \delta[$ dans

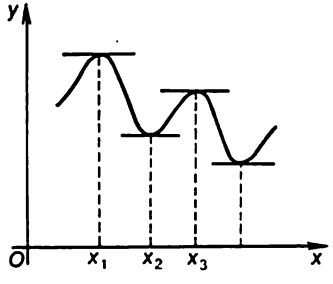


Fig. 77

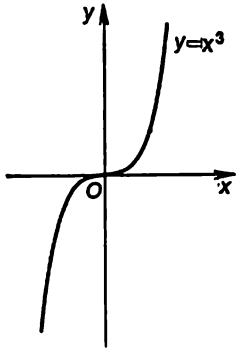


Fig. 78

lequel $f(x_0)$ est la plus grande ou la plus petite des valeurs prises sur cet intervalle. Le théorème de Fermat nous dit alors que $f'(x_0) = 0$. ■

Le théorème 6.8 admet l'interprétation géométrique suivante. Si une fonction présente des extrémums locaux en x_1, x_2 et x_3 et si sa courbe représentative admet des tangentes aux points d'abscisses x_1, x_2 et x_3, alors ces tangentes sont parallèles à Ox (fig. 77).

Ces points sont parfois appelés *stationnaires*; nous les appellerons *points d'éventuel extrémum*. En un point d'éventuel extrémum x_0, c'est-à-dire tel que $f'(x_0) = 0$, la fonction $f(x)$ ne présente pas nécessairement un maximum ou un minimum local. Par exemple, si $f(x) = x^3$, alors $f'(x) = 3x^2 = 0$ en $x = 0$, pourtant en $x = 0$ cette fonction ne présente pas d'extrémum local (fig. 78). Etablissons une condition suffisante d'existence d'un extrémum local.

Théorème 6.9 (condition suffisante d'extrémum local). *Soit $f(x)$ une fonction dérivable dans un δ-voisinage d'un point x_0. Si $f'(x) > 0$ (resp. $f'(x) < 0$) pour tout $x \in]x_0 - \delta, x_0[$ et $f'(x) < 0$ (resp. $f'(x) > 0$) pour tout $x \in]x_0, x_0 + \delta[$, alors $f(x)$ présente un maximum* (resp. *un minimum*) *local en x_0; si $f'(x)$ ne change pas de signe sur ce δ-voisinage tout entier, $f(x)$ ne présente pas d'extrémum local en x_0.*

En d'autres termes, lorsqu'on passe par x_0, la fonction $f(x)$ présente un maximum local en x_0 si $f'(x)$ passe du signe + au signe —; un minimum local, si $f'(x)$ passe du signe — au signe +; ne présente pas d'extrémum si $f'(x)$ garde son signe.

Démonstration. Supposons que lorsqu'on passe par x_0, la dérivée $f'(x)$ passe du signe + au signe —, et soit $x \in]x_0 - \delta, x_0[$. Appliquons la formule de Lagrange à la fonction $f(x)$ sur le segment $[x, x_0]$. On obtient

$$f(x_0) - f(x) = f'(c)(x_0 - x), \quad c \in]x, x_0[.$$

Puisque $f'(x) > 0$ sur $]x_0 - \delta, x_0[$, on a $f'(c) > 0$ et de plus $x_0 - x > 0$, donc

$$f(x_0) - f(x) > 0 \text{ ou } f(x_0) > f(x). \tag{1}$$

Traitons maintenant le cas où $x \in]x_0, x_0 + \delta[$. Appliquons la formule de Lagrange à la fonction $f(x)$ sur le segment $[x_0, x]$. On obtient

$$f(x) - f(x_0) = f'(c)(x - x_0), \quad c \in]x_0, x[.$$

Comme $f'(x) < 0$ sur $]x_0, x_0 + \delta[$, on a $f'(c) < 0$ et de plus $x - x_0 > 0$, donc

$$f(x) - f(x_0) < 0 \text{ ou } f(x_0) > f(x). \tag{2}$$

Des inégalités (1) et (2) il s'ensuit que dans le δ-voisinage du point x_0, on a $f(x) < f(x_0)$ pour $x \neq x_0$, ce qui exprime que la fonction $f(x)$ présente un maximum local en x_0.

On traite de même le cas où $f'(x)$ passe du signe — au signe +. Reste le cas où $f'(x)$ ne change pas de signe. Supposons que $f'(x) > 0$ dans un voisinage $]x_0 - \delta, x_0 + \delta[$; alors d'après le théorème 6.7 la fonction $f(x)$ est croissante sur ce voisinage, c'est-à-dire que $f(x) < f(x_0)$ pour tout $x < x_0$ et $f(x) > f(x_0)$ pour tout $x > x_0$. Ce qui exprime que $f(x)$ ne présente pas d'extrémum en x_0. ■

Remarque. Le théorème 6.9 reste valable si la fonction $f(x)$ n'est pas dérivable mais seulement continue en x_0. Par exemple, la fonction $f(x) = |x|$ est continue en $x = 0$ mais pas dérivable.

A titre d'exemple cherchons les points en lesquels la fonction $f(x) = x^3 - 3x$ présente un extrémum local. Calculons la dérivée $f'(x) = 3x^2 - 3 = 3(x^2 - 1)$. La résolution de l'équation

x	$-\infty$	-1		1	$+\infty$
$f'(x)$	$+$	0	$-$	0	$+$
$f(x)$	↗	2	↘	2	↗

Fig. 79

$3(x^2 - 1) = 0$ nous donne deux points d'éventuel extrémum: $x_1 = -1$ et $x_2 = 1$. Il est plus commode de poursuivre notre étude sur le tableau de variation de la figure 79. On voit que la fonction $f(x)$ présente un maximum local égal à 2 en $x_1 = -1$ et un minimum local égal à -2 en $x_2 = 1$.

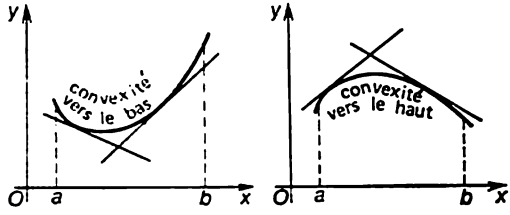


Fig. 80

Par ailleurs, $f(x)$ est strictement croissante sur $]-\infty, -1[$ et $]1, +\infty[$ et strictement décroissante sur $]-1, 1[$.

3. Sens de convexité et point d'inflexion de la courbe représentative d'une fonction. Soit $y = f(x)$ une fonction dérivable sur un intervalle $]a, b[$. La courbe représentative de la fonction $y = f(x)$ admet alors une tangente en tout point $M(x; f(x))$, $x \in]a, b[$, et de plus cette tangente n'est pas parallèle à l'axe Oy, puisque son coefficient directeur qui est égal à $f'(x)$ est fini.

Définition 1. On dira que la courbe représentative d'une fonction $y = f(x)$ a sa *convexité tournée vers le bas* (resp. *vers le haut*) sur un intervalle $]a, b[$ si elle est située au-dessus (resp. au-dessous) de toute tangente sur $]a, b[$ (fig. 80).

Théorème 6.10. *Si une fonction $y = f(x)$ admet sur un intervalle $]a, b[$ une dérivée seconde $f''(x) \geqslant 0$ (resp. $f''(x) \leqslant 0$), sa courbe représentative est convexe vers le bas* (resp. *vers le haut*) *sur $]a, b[$.*

Démonstration. Considérons pour fixer les idées le cas où $f''(x) \geqslant 0$ sur $]a, b[$. Soit c un point quelconque de $]a, b[$ (fig. 81). On demande de prouver que la courbe $y = f(x)$ est située au-dessus de la tangente passant par le point $M(c; f(c))$.

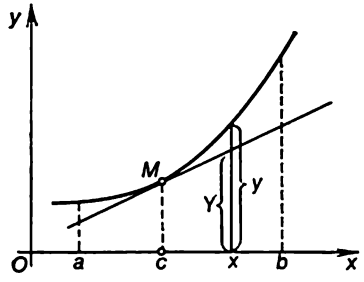


Fig. 81

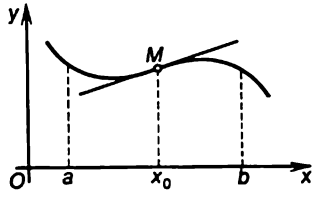


Fig. 82

Ecrivons l'équation de cette tangente en désignant par Y l'ordonnée courante de ses points: $Y - f(c) = f'(c)(x - c)$ ou

$$Y = f(c) + f'(c)(x - c). \tag{3}$$

Développons la fonction $f(x)$ au voisinage du point c à l'aide de la formule de Taylor pour $n = 1$. On obtient

$$y = f(x) = f(c) + \frac{f'(c)}{1!}(x - c) + \frac{f''(\xi)}{2!}(x - c)^2, \qquad \xi \in]c, x[. \tag{4}$$

La formule (4) est valable pour tout $x \in]a, b[$.

En soustrayant (3) de (4), on obtient

$$y - Y = \frac{f''(\xi)}{2!}(x - c)^2. \tag{5}$$

Puisque $f''(x) \geqslant 0$ sur $]a, b[$ par hypothèse, le second membre de (5) est positif, c'est-à-dire que $y - Y \geqslant 0$ pour tout $x \in]a, b[$, ou $y \geqslant Y$. La dernière inégalité exprime que la courbe $y = f(x)$ est située au-dessus de la tangente (3) sur $]a, b[$.

Le théorème se prouve par analogie pour $f''(x) \leqslant 0$. ■

Définition 2. On dit qu'un point $M(x_0; f(x_0))$ est un *point d'inflexion* d'une courbe $y = f(x)$ si cette courbe admet une tangente en M et s'il existe un voisinage du point x_0 dans lequel elle a sa convexité tournée dans des sens différents de part et d'autre du point x_0.

Il est évident que la tangente traverse la courbe au point d'inflexion, puisque la courbe est située au-dessus de cette tangente d'un côté de ce point et au-dessous, de l'autre côté, autrement dit au voisinage du point d'inflexion la courbe traverse la tangente en s'« infléchissant ». D'où l'origine du nom « point d'inflexion » (fig. 82).

Théorème 6.11 (condition nécessaire d'inflexion). *Supposons qu'une courbe* $y = f(x)$ *présente une inflexion au point* $M(x_0; f(x_0))$ *et que la fonction* $y = f(x)$ *admet une dérivée seconde continue en* x_0. *Alors* $f''(x_0) = 0$.

Démonstration. Supposons par absurde que $f''(x_0) \neq 0$. Puisque la dérivée seconde est continue, le théorème 4.8 de stabilité du signe d'une fonction continue affirme l'existence d'un voisinage du point x_0 dans lequel $f''(x) < 0$ (ou $f''(x) > 0$), donc d'après le théorème 6.10 la courbe $y = f(x)$ a sa convexité tournée dans un sens bien défini dans ce voisinage. Ceci exprime que $M(x_0; f(x_0))$ (fig. 82) n'est pas un point d'inflexion. Cette contradiction prouve le théorème. ■

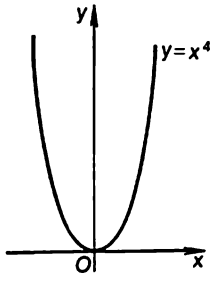


Fig. 83

Signalons que $f''(x_0) = 0$ ne signifie pas nécessairement que $M(x_0; f(x_0))$ est un point d'inflexion. Par exemple, la courbe d'équation $f(x) = x^4$ ne présente pas d'inflexion en (0; 0) bien que $f''(x) = 12x^2 = 0$ pour $x = 0$ (fig. 83). Donc la condition de nullité de la dérivée seconde n'est qu'une condition nécessaire d'inflexion. Les points $M(x_0; f(x_0))$ de la courbe en lesquels $f''(x_0) = 0$ s'appellent *points critiques*. Une étude supplémentaire est nécessaire pour dire si un point critique est point d'inflexion. Etablissons à cet effet une condition suffisante d'inflexion.

Théorème 6.12 (condition suffisante d'inflexion). *Supposons qu'une fonction* $y = f(x)$ *admet une dérivée seconde dans un voisinage d'un point* x_0. *Si dans ce voisinage* $f''(x)$ *a des signes différents à gauche et à droite du point* x_0, *la courbe d'équation* $y = f(x)$ *présente une inflexion au point* $M(x_0; f(x_0))$.

Démonstration. Du fait que $f''(x)$ a des signes différents à gauche et à droite du point x_0, il s'ensuit en vertu du théorème 6.10 que la convexité de la courbe est tournée dans des sens différents à gauche et à droite de x_0. Ce qui exprime que $M(x_0; f(x_0))$ est un point d'inflexion. ■

Remarque. Le théorème reste en vigueur si $f(x)$ admet une dérivée seconde dans un voisinage du point x_0, excepté x_0, et si la courbe représentative de $f(x)$ admet une tangente en M. Dans ces conditions, si dans ce voisinage $f''(x)$ a des signes différents à gauche et à droite de x_0, la courbe $y = f(x)$ présente une inflexion en $M(x_0; f(x_0))$. La démonstration de ce fait est calquée sur celle du théorème.

Considérons un exemple: $f(x) = x^{1/3}$. Cette fonction admet une dérivée infinie en $x = 0$ et la tangente à sa courbe représentative au point O (0; 0) est confondue avec l'axe Oy. La dérivée seconde n'existe pas en $x = 0$. Néanmoins la courbe $y = x^{1/3}$ présente une inflexion en O (0; 0) puisque la dérivée seconde $f''(x) = 2/(9x^{5/3})$ a des signes différents à gauche et à droite du point $x = 0$ (fig. 84).

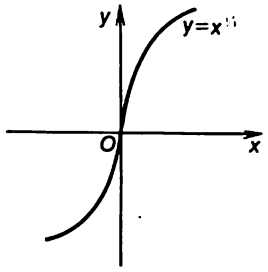


Fig. 84

Ainsi la dérivée seconde sert à étudier le sens de convexité et les points d'inflexion de la courbe représentative d'une fonction. Poursuivons à titre d'exemple l'étude de la fonction $f(x) = x^3 - 3x$ qui a été entamée au n° 2. La dérivée seconde est $f''(x) = 6x$. Donc le point O (0; 0) est un point critique. Sur le tableau

x	$-\infty$		-1		0		1		$+\infty$
$f'(x)$	$+$		0	$-$		$-$	0	$+$	
$f''(x)$	$-$			$-$	0	$+$		$+$	

Fig. 85

(fig. 85) où l'on étudie le signe de $f''(x)$, on voit que $f''(x) < 0$ à gauche de 0 (ce qui exprime que la courbe est convexe vers le haut) et > 0 à droite de 0 (la courbe est convexe vers le bas). La courbe représentative de la fonction $f(x) = x^3 - 3x$ est tracée sur la figure 86.

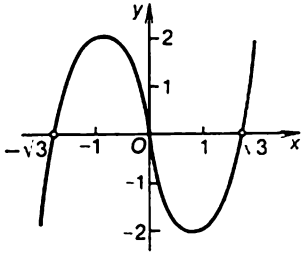


Fig. 86

Montrons maintenant que la partie d'ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ située dans le demi-plan supérieur ($y \geqslant 0$) est convexe vers le haut sur l'intervalle $]-a, a[$. En effet, on a $y = \frac{b}{a}\sqrt{a^2 - x^2}$. D'autre part, $y' = -\frac{b}{a}\frac{x}{\sqrt{a^2 - x^2}}$, $y'' = -\frac{ba}{(a^2 - x^2)^{3/2}}$. On remarque que la dérivée seconde est strictement négative sur l'intervalle $]-a, a[$. Donc la courbe envisagée est convexe vers le haut sur l'intervalle $]-a, a[$ (cf. fig. 33).

On démontre de même (ce qu'on laisse au soin du lecteur) que la portion d'hyperbole $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ comprise dans le demi-plan supérieur est convexe vers le haut sur les intervalles $]a, +\infty[$ et $]-\infty, -a[$.

4. Asymptotes de la courbe représentative d'une fonction. Lorsqu'on étudie le comportement d'une fonction à l'infini, c'est-à-dire pour $x \to \pm\infty$ ou au voisinage des points de discontinuité de deuxième espèce, on constate souvent que sa courbe représentative se rapproche aussi près que l'on veut d'une ou de plusieurs droites appelées *asymptotes* *).

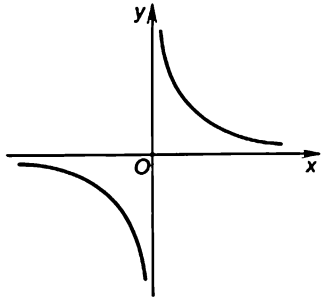


Fig. 87

Les asymptotes peuvent être verticales, horizontales ou obliques.

Définition 1. On dit qu'une droite $x = x_0$ est une *asymptote verticale* d'une courbe d'équation $y = f(x)$ si l'une au moins des limites $\lim\limits_{x \to x_0+} f(x)$ ou $\lim\limits_{x \to x_0-} f(x)$ est égale à $+\infty$ ou $-\infty$.

Dans ce cas la distance du point courant $M(x; f(x))$ à la droite $x = x_0$ est égale à $d = \sqrt{(x - x_0)^2 + (f(x) - f(x))^2} = |x - x_0|$ et par suite $d \to 0$ lorsque $x \to x_0$.

Par exemple, la courbe représentative de la fonction $y = f(x) = 1/x$ (fig. 87) admet l'asymptote verticale $x = 0$, puisque $f(x) \to +\infty$ lorsque $x \to 0+$ et $f(x) \to -\infty$ lorsque $x \to 0-$.

Définition 2. Une droite $y = A$ est une *asymptote horizontale* de la courbe $y = f(x)$ lorsque $x \to \pm\infty$ si $\lim\limits_{x \to \pm\infty} f(x) = A$.

Dans ce cas la distance du point courant $M(x; f(x))$ à la droite $y = A$ est égale à $d = \sqrt{(x - x)^2 + (f(x) - A)^2} = |f(x) - A|$ et par suite $d \to 0$ pour $x \to \infty$, puisque $\lim\limits_{x \to \infty} |f(x) - A| = 0$.

Par exemple, la courbe $y = 1/x$ admet l'asymptote horizontale $y = 0$ pour $x \to \pm\infty$, puisque $1/x \to 0$ pour $x \to \pm\infty$.

Définition 3. On dit qu'une droite $y = kx + b$ $(k \neq 0)$ est une *asymptote oblique* de la courbe $y = f(x)$ lorsque $x \to \pm\infty$ si

*) La notion d'asymptote n'est pas nouvelle, nous l'avons déjà rencontrée en géométrie analytique en étudiant l'hyperbole (cf. chap. III, § 6, n° 2).

la fonction $f(x)$ peut être mise sous la forme

$$f(x) = kx + b + \alpha(x), \tag{6}$$

où $\alpha(x) \to 0$ lorsque $x \to \pm\infty$.

Voyons la signification géométrique de l'asymptote oblique. Pour fixer les idées on traitera le cas où $x \to +\infty$ (le cas $x \to -\infty$ se traitant de façon analogue).

Soit $M(x; y)$ un point de la courbe $y = f(x)$ et soit $\tilde{y} = kx + b$ une asymptote oblique lorsque $x \to +\infty$. Désignons par $N(x; \tilde{y})$ le point courant de l'asymptote, $\tilde{y}$ étant l'ordonnée courante (fig. 88). Alors $MN = |y - \tilde{y}| = |f(x) - (kx + b)| = |\alpha(x)| \to 0$ lorsque $x \to +\infty$. Abaissons la perpendiculaire MP sur l'asymptote. La distance d de M à l'asymptote est égale à $MP = MN \cos\alpha$, où α est l'angle de l'asymptote et de l'axe Ox, donc $\lim\limits_{x \to +\infty} d = 0$.

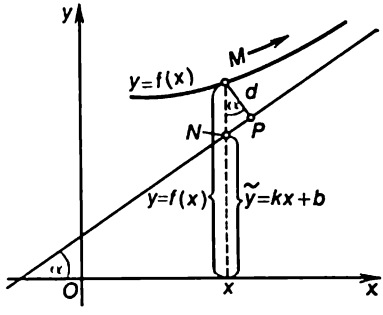


Fig. 88

Donc la distance de $M(x; y)$ tend vers 0 lorsque $x \to +\infty$, autrement dit la courbe $y = f(x)$ tend indéfiniment vers l'asymptote lorsque $x \to +\infty$.

Etablissons l'équation de l'asymptote oblique, c'est-à-dire calculons les nombres k et b de l'équation de l'asymptote. En divisant l'égalité (6) par x et en faisant tendre x vers ∞, on obtient

$$\lim_{x \to +\infty} \frac{f(x)}{x} = \lim_{x \to +\infty} \left[k + \frac{b}{x} + \frac{\alpha(x)}{x}\right] = k.$$

Ainsi

$$k = \lim_{x \to +\infty} \frac{f(x)}{x}. \tag{7}$$

Par ailleurs, la relation (6) nous donne: $\lim\limits_{x \to +\infty} [f(x) - kx] = \lim\limits_{x \to +\infty} [b + \alpha(x)] = b$. Donc

$$b = \lim_{x \to +\infty} [f(x) - kx]. \tag{8}$$

On a ainsi prouvé que si la droite $\tilde{y} = kx + b$ est une asymptote oblique, les nombres k et b sont donnés par les formules (7) et (8). Réciproquement, si les limites (7) et (8) existent toutes les

deux, et de plus $k \neq 0$, alors la droite $\tilde{y} = kx + b$ est une asymptote oblique de la courbe $y = f(x)$ lorsque $x \to +\infty$.

En effet, en posant $\alpha(x) = f(x) - kx - b$ et en se servant de l'égalité (8), on trouve que $\lim\limits_{x \to +\infty} \alpha(x) = 0$. On a donc l'égalité (6): $f(x) = kx + b + \alpha(x)$, où $\lim\limits_{x \to +\infty} \alpha(x) = 0$, c'est-à-dire que la droite $\tilde{y} = kx + b$ est une asymptote oblique lorsque $x \to +\infty$.

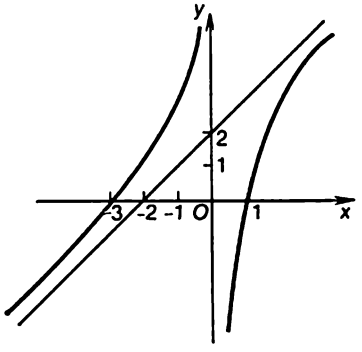


Fig. 89

Les asymptotes doivent être cherchées dans l'ordre suivant:
1) les asymptotes verticales;
2) les asymptotes horizontales;
3) les asymptotes obliques.

Exemple 1. Trouver les asymptotes de la courbe $y = \dfrac{x^2+2x-3}{x}$.

S o l u t i o n. 1) Asymptotes verticales. Le point $x = 0$ est un point de discontinuité de deuxième espèce, de plus $y \to +\infty$ lorsque $x \to 0-$ et $y \to -\infty$ lorsque $x \to 0+$. Donc l'axe $x = 0$ est une asymptote verticale.

2) Asymptotes horizontales:

$$\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{x^2+2x-3}{x} = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left(x + 2 - \frac{3}{x}\right) = +\infty\,(-\infty),$$

donc il n'existe pas d'asymptote horizontale.

3) Asymptotes obliques:

$$k = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{f(x)}{x} = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left(1 + \frac{2}{x} - \frac{3}{x^2}\right) = 1,$$

$$b = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} [f(x) - kx] = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left[\frac{x^2+2x-3}{x} - x\right] =$$

$$= \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left(\frac{2x-3}{x}\right) = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left(2 - \frac{3}{x}\right) = 2.$$

Donc la droite $y = x + 2$ est une asymptote oblique lorsque $x \to \pm\infty$.

La courbe correspondante est représentée sur la figure 89.

Exemple 2. Montrer que l'hyperbole $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ a pour asymptotes obliques les droites $y = \pm \dfrac{b}{a} x$.

Solution. Comme $y = \pm \frac{b}{a} \sqrt{x^2 - a^2}$, alors

$$\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{f(x)}{x} = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left[\pm \frac{b}{a} \sqrt{1 - \left(\frac{a}{x}\right)^2} \right] = \pm \frac{b}{a};$$

$$\lim_{\substack{x \to +\infty \\ (x \to -\infty)}} [f(x) - kx] = \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left[\pm \frac{b}{a} \sqrt{x^2 - a^2} \mp \frac{b}{a} x \right] =$$

$$= \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \left[\pm \frac{b}{a} (\sqrt{x^2 - a^2} - x) \right] = \pm \frac{b}{a} \lim_{\substack{x \to +\infty \\ (x \to -\infty)}} \frac{-a^2}{\sqrt{x^2 - a^2} + x^2} = 0.$$

Donc les droites $y = \pm \frac{b}{a} x$ sont des asymptotes obliques de cette hyperbole lorsque $x \to \pm\infty$.

5. Plan d'étude d'une fonction. L'étude du comportement d'une fonction avec le tracé de sa courbe comprend les étapes suivantes.

a) Détermination de l'ensemble de définition et étude de la continuité.

b) Détermination des symétries ou des périodes: ceci permet de limiter l'ensemble sur lequel est étudiée la fonction.

c) Détermination du sens de variation soit par une étude directe, soit par une étude du sens de la dérivée.

d) Etude de la fonction aux bornes des intervalles de définition et de continuité.

e) Tableau de variation.

f) Etude des branches infinies: recherche des asymptotes et des directions asymptotiques.

g) Recherche des points remarquables: points d'éventuels extrémums, points critiques, points d'arrêt, points anguleux, points d'inflexion, points d'intersection avec les axes de coordonnées, etc.

Exemple. $y = \frac{x^2 + 1}{x - 1}$.

a) y est définie et continue pour tout $x \neq 1$.

b) Pas de symétrie (puisque $y(x) \neq \pm y(-x)$), ni de période.

c) $y'(x) = \frac{2x(x-1) - (x^2+1)}{(x-1)^2} = \frac{2x^2 - 2x - x^2 - 1}{(x-1)^2} = \frac{x^2 - 2x - 1}{(x-1)^2}$.

Cette dérivée est du signe de $x^2 - 2x - 1$ dont les racines sont $x = 1 + \sqrt{2}$ et $x = 1 - \sqrt{2}$. On a

$$y(1 + \sqrt{2}) = 2 + 2\sqrt{2}, \quad y(1 - \sqrt{2}) = 2 - 2\sqrt{2}.$$

d) Lorsque $x \to \pm\infty$, $y \to \pm\infty$.

e) Tableau de variation (cf. fig. 90).

Le tableau met en évidence l'existence d'une asymptote verticale d'équation $x = 1$.

f) Branches infinies lorsque $x \to \pm\infty$:

$$k = \lim_{\substack{x\to+\infty \\ (x\to-\infty)}} \frac{f(x)}{x} = \lim_{\substack{x\to+\infty \\ (x\to-\infty)}} \frac{x^2+1}{x^2-x} = \lim_{\substack{x\to+\infty \\ (x\to-\infty)}} \frac{1+1/x^2}{1-1/x} = 1,$$

donc une direction asymptotique de coefficient directeur 1 ;

$$b = \lim_{\substack{x\to+\infty \\ (x\to-\infty)}} [f(x)-kx] = \lim_{\substack{x\to+\infty \\ (x\to-\infty)}} \left[\frac{x^2+1}{x-1} - x\right] =$$

$$= \lim_{\substack{x\to+\infty \\ (x\to-\infty)}} \frac{1+x}{x-1} = \lim_{\substack{x\to+\infty \\ (x\to-\infty)}} \frac{1+1/x}{1-1/x} = 1,$$

donc une asymptote oblique d'équation $y = x + 1$.

x	$-\infty$		$1-\sqrt{2}$		1		$1+\sqrt{2}$		$+\infty$
y'		$+$	0	$-$	‖	$-$	0	$+$	
y		↗	$2-\sqrt{2}$	↘	$-\infty$ ‖ $+\infty$	↘	$2+2\sqrt{2}$	↗	

Fig. 90

g) Points remarquables. La fonction y présente un maximum local au point $(1-\sqrt{2},\ 2-2\sqrt{2})$ et un minimum local au point $(1+\sqrt{2},\ 2+2\sqrt{2})$.

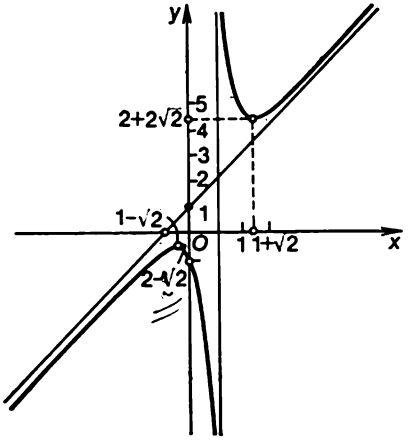


Fig. 91

Calculons la dérivée seconde pour déterminer les points critiques. On a

$$y'' = \frac{(2x-2)(x-1)^2 - 2(x-1)(x^2-2x-1)}{(x-1)^4} = \frac{4}{(x-1)^3}.$$

Comme y'' n'est jamais nulle, la courbe y ne présente pas de points critiques. Etudions le signe de y'' pour déterminer le sens de convexité de la courbe y.

x	$-\infty$		$1-\sqrt{2}$		1		$1+\sqrt{2}$		$+\infty$
y'	$+$		0	$-$	‖	$-$	0	$+$	
y''	$-$			$-$	‖	$+$		$+$	

Donc la courbe y est convexe vers le haut sur $]-\infty, 1[$ et vers le bas sur $]1, +\infty[$.

La courbe représentative de la fonction $y = \frac{x^2+1}{x-1}$ est tracée sur la figure 91 (les figures 90 et 91 nécessitent une correction : pour $2-\sqrt{2}$ veuillez lire $2-2\sqrt{2}$).

§ 5. Interpolation des fonctions

L'interpolation est utilisée dans de nombreux problèmes aussi bien théoriques que pratiques nécessitant des calculs.

1. Position du problème. Soient données sur un segment $[a, b]$ les valeurs d'une fonction $y = f(x)$ aux points $a \leqslant x_0 < x_1 < x_2 < \ldots < x_n \leqslant b$:

$$f(x_0) = y_0,\ f(x_1) = y_1,\ f(x_2) = y_2,\ \ldots,\ f(x_n) = y_n.$$

On demande de trouver un polynôme de degré $\leqslant n$:

$$P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n, \qquad (1)$$

prenant en $x_0, x_1, \ldots, x_n$ les mêmes valeurs que la fonction $f(x)$, c'est-à-dire que

$$P_n(x_i) = f(x_i) = y_i,\ i = 0, 1, 2, \ldots, n. \qquad (2)$$

Autrement dit, on demande de trouver un polynôme de la forme (1) qui soit une approximation de la fonction $y = f(x)$ sur le segment $[a, b]$. Le problème posé s'appelle *problème d'interpolation*, le polynôme (1), *polynôme d'interpolation*, et les points $x_0, x_1, \ldots, x_n$, *nœuds d'interpolation*. La solution de ce problème permet de trouver les valeurs approchées de la fonction $f(x)$ aux points x compris entre les nœuds. Ceci est important lorsque la fonction n'est donnée qu'aux points $x_0, x_1, \ldots, x_n$ et qu'il faut déterminer ses valeurs dans les intervalles de ces points et aussi lorsque la fonction $f(x)$ est donnée par une formule sur le segment $[a, b]$ tout entier et que le calcul de ses valeurs par cette formule est trop laborieux.

Montrons qu'il existe toujours un polynôme d'interpolation (1) et un seul satisfaisant les conditions (2). Bornons-nous pour

simplifier au cas $n = 2$, c'est-à-dire à un trinôme du second degré

$$P_2(x) = a_0x^2 + a_1x + a_2. \qquad (3)$$

En substituant x_0, x_1 et x_2 à x dans l'équation (3) et comme le polynôme (3) prend respectivement les valeurs y_0, y_1 et y_2 en ces points, on obtient le système de trois équations du premier degré à trois coefficients inconnus a_0, a_1 et a_2 :

$$\begin{cases} a_0x_0^2 + a_1x_0 + a_2 = y_0, \\ a_0x_1^2 + a_1x_1 + a_2 = y_1, \\ a_0x_2^2 + a_1x_2 + a_2 = y_2. \end{cases}$$

Les nombres x_0, x_1 et x_2 étant distincts, le déterminant de ce système est non nul :

$$\begin{vmatrix} x_0^2 & x_0 & 1 \\ x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \end{vmatrix} = (x_1 - x_0)(x_2 - x_0)(x_1 - x_2) \neq 0.$$

Donc ce système admet une solution et une seule (cf. chap. X, § 3), ce que nous voulions. Géométriquement cela signifie que par trois points $M_0(x_0; y_0)$, $M_1(x_1; y_1)$ et $M_2(x_2; y_2)$ il passe une seule courbe d'équation (3). Donc il existe toujours un polynôme d'interpolation (1) et un seul. On étudiera plus bas les diverses formes du polynôme d'interpolation.

2. Formule d'interpolation de Lagrange. Déterminons les coefficients du polynôme d'interpolation (1). En portant ce polynôme dans le système (2), on obtient un système de $n + 1$ équations du premier degré à $n + 1$ coefficients inconnus a_0, a_1, ... , a_n :

$$\begin{cases} a_0 + a_1x_0 + \ldots + a_nx_0^n = y_0, \\ a_0 + a_1x_1 + \ldots + a_nx_1^n = y_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_0 + a_1x_n + \ldots + a_nx_n^n = y_n. \end{cases}$$

En portant les coefficients a_0, a_1, . . . , a_n obtenus par la résolution de ce système dans l'égalité (1), on trouve le polynôme d'interpolation cherché. Mais la résolution du système donne généralement lieu à des calculs inextricables. Aussi cherchera-t-on le polynôme (1) sous la forme

$$\begin{aligned} P_n(x) = {} & a_0(x - x_1)(x - x_2) \ldots (x - x_n) + \\ & + a_1(x - x_0)(x - x_2) \ldots (x - x_n) + \\ & + a_2(x - x_0)(x - x_1)(x - x_3) \ldots (x - x_n) + \ldots \\ & \ldots + a_n(x - x_0)(x - x_1)(x - x_2) \ldots (x - x_{n-1}). \end{aligned} \qquad (4)$$

En faisant $x = x_0$ dans (4) et en tenant compte de la condition (2), on obtient

$$y_0 = a_0 (x_0 - x_1)(x_0 - x_2) \ldots (x_0 - x_n),$$

d'où

$$a_0 = \frac{y_0}{(x_0 - x_1)(x_0 - x_2) \ldots (x_0 - x_n)}.$$

En faisant ensuite $x = x_1$ dans (4), on trouve

$$y_1 = a_1 (x_1 - x_0)(x_1 - x_2) \ldots (x_1 - x_n),$$

d'où

$$a_1 = \frac{y_1}{(x_1 - x_0)(x_1 - x_2) \ldots (x_1 - x_n)}.$$

De façon analogue

$$a_2 = \frac{y_2}{(x_2 - x_0)(x_2 - x_1)(x_2 - x_3) \ldots (x_2 - x_n)},$$

$$\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$$

$$a_n = \frac{y_n}{(x_n - x_0)(x_n - x_1)(x_n - x_2) \ldots (x_n - x_{n-1})}.$$

La substitution des valeurs des coefficients $a_0, a_1, \ldots, a_n$ dans la formule (4) nous donne le polynôme d'interpolation cherché:

$$P_n(x) = \frac{(x - x_1)(x - x_2) \ldots (x - x_n)}{(x_0 - x_1)(x_0 - x_2) \ldots (x_0 - x_n)} y_0 + \frac{(x - x_0)(x - x_2) \ldots (x - x_n)}{(x_1 - x_0)(x_1 - x_2) \ldots (x_1 - x_n)} y_1 + \ldots + \frac{(x - x_0)(x - x_1) \ldots (x - x_{n-1})}{(x_n - x_0)(x_n - x_1) \ldots (x_n - x_{n-1})} y_n. \quad (5)$$

La formule (5) s'appelle *formule d'interpolation de Lagrange.*

Exemple. Les valeurs empiriques prises par une fonction $f(x)$ aux points $x_0 = 1$, $x_1 = 3$ et $x_2 = 5$ sont respectivement $y_0 = 2$, $y_1 = 1$ et $y_2 = 8$. Trouver le polynôme du second degré $P_2(x)$ qui approche la fonction $f(x)$.

Solution. La formule (5) nous donne

$$P_2(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} y_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} y_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} y_2 = \frac{(x - 3)(x - 5)}{(1 - 3)(1 - 5)} 2 + \frac{(x - 1)(x - 5)}{(3 - 1)(3 - 5)} 1 + \frac{(x - 1)(x - 3)}{(5 - 1)(5 - 3)} 8;$$

$$P_2(x) = x^2 - (9/2)\, x + 11/2.$$

3. Formule d'interpolation de Newton. Considérons le cas particulier où la différence h de deux nœuds voisins est constante:

$x_i - x_{i-1} = h$. Introduisons les notations suivantes:

$$\Delta y_0 = y_1 - y_0,\ \Delta y_1 = y_2 - y_1,\ \Delta y_2 = y_3 - y_2,\ \ldots,$$
$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0,\ \Delta^2 y_1 = \Delta y_2 - \Delta y_1,\ \Delta^2 y_2 = \Delta y_3 - \Delta y_2,\ \ldots,$$
$$\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0,\ \Delta^3 y_1 = \Delta^2 y_2 - \Delta^2 y_1,\ \ldots,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\Delta^n y_0 = \Delta^{n-1} y_1 - \Delta^{n-1} y_0,\ \Delta^n y_1 = \Delta^{n-1} y_2 - \Delta^{n-1} y_1,\ \ldots,$$

appelées *différences du premier, du second, . . ., du n-ème ordre.*

Trouver le polynôme d'interpolation de degré n qui prend les valeurs $y_0, y_1, y_2, \ldots, y_n$ respectivement aux points $x_0, x_1 = x_0 + h, x_2 = x_0 + 2h, \ldots, x_n = x_0 + nh$. Trouvons d'abord le polynôme du premier degré prenant les valeurs y_0 et y_1 aux points x_0 et $x_1 = x_0 + h$. En substituant $x_1 = x_0 + h$ à x_1 dans (5), on obtient

$$P_1(x) = y_0 + \Delta y_0 \frac{x - x_0}{h}.$$

De façon analogue,

$$P_2(x) = y_0 + \Delta y_0 \frac{x - x_0}{h} + \frac{\Delta^2 y_0}{2!\,h^2}(x - x_0)(x - x_1),$$

$$P_3(x) = y_0 + \Delta y_0 \frac{x - x_0}{h} + \frac{\Delta^2 y_0}{2!\,h^2}(x - x_0)(x - x_1) +$$
$$+ \frac{\Delta^3 y_0}{3!\,h^3}(x - x_0)(x - x_1)(x - x_2),$$

et, d'une façon générale,

$$P_n(x) = y_0 + \Delta y_0 \frac{x - x_0}{h} + \frac{\Delta^2 y_0}{2!\,h^2}(x - x_0)(x - x_1) +$$
$$+ \frac{\Delta^3 y_0}{3!\,h^3}(x - x_0)(x - x_1)(x - x_2) + \ldots$$
$$\ldots + \frac{\Delta^n y_0}{n!\,h^n}(x - x_0)(x - x_1)(x - x_2)\ldots(x - x_{n-1}). \qquad (6)$$

La formule (6) qui s'appelle *formule d'interpolation de Newton* définit le polynôme cherché.

Le problème d'interpolation admet une seule solution, donc les formules de Lagrange et de Newton pour des valeurs données x_i et y_i sont identiques et ne diffèrent que par le groupement des termes. Dans la pratique la formule de Newton est plus commode pour la raison suivante: l'adjonction de nouveaux nœuds d'interpolation implique pour la formule de Lagrange le calcul de tous les coefficients, pour la formule de Newton, le calcul des nouveaux termes seulement, les anciens restant les mêmes.

Il existe d'autres formules d'interpolation dont la plus courante est l'*interpolation d'Hermite.* Le problème se pose dans les termes suivants: étant donnés n nœuds, n valeurs d'une fonction $f(x)$

et n valeurs de sa dérivée $f'(x)$ en ces nœuds, on demande de trouver un polynôme de degré $\leqslant 2n-1$ tel que

$$P(x_i) = f(x_i) = y_i,\ P'(x_i) = f'(x_i) = y'_i,\ i = 1, 2, \ldots, n.$$

On ne s'attardera pas sur la résolution de ce problème, on remarquera seulement que si les x_i sont tous distincts, il existe une seule solution que l'on trouve comme plus haut.

4. **Reste d'une interpolation.** Pour apprécier le degré de proximité du polynôme d'interpolation $P_n(x)$ de la fonction $f(x)$ il faut étudier leur différence

$$R(x) = f(x) - P_n(x)$$

qui s'appelle *reste de l'interpolation.*

Supposons que $f(x)$ est $n+1$ fois dérivable sur $[a, b]$. Alors

$$R^{(n+1)}(x) = f^{(n+1)}(x),\quad x \in [a, b], \tag{7}$$

puisque $P_n^{(n+1)}(x) \equiv 0$. Soient x un nombre fixe quelconque de $[a, b]$ non confondu avec un nœud d'interpolation, $t \in [a, b]$ une variable. Posons

$$\omega(x) = (x-x_0)(x-x_1)(x-x_2)\ldots(x-x_n)$$

et considérons sur $[a, b]$ la fonction auxiliaire

$$F(t) = R(t) - \frac{R(x)}{\omega(x)}\omega(t),\ t \in [a, b].$$

La fonction $F(t)$ est visiblement $n+1$ fois dérivable sur $[a, b]$ et de plus, en vertu de (7) et du fait que $\omega^{(n+1)}(t) = (n+1)!$, on obtient

$$F^{(n+1)}(t) = f^{(n+1)}(t) - (n+1)!\,\frac{R(x)}{\omega(x)}. \tag{8}$$

Par ailleurs, la fonction $F(t)$ s'annule en $n+2$ points : $x_0, x_1, \ldots, x_n$ et x (fixe). Donc sa dérivée première s'annule en vertu du théorème de Rolle au moins en $n+1$ points sur $[a, b]$, la dérivée seconde s'annule au moins en n points sur $[a, b]$, etc. Par récurrence on établit que la dérivée $F^{(n+1)}(t)$ s'annule au moins une fois à l'interieur de $[a, b]$. Il existe donc un point $\xi \in]a, b[$ tel que

$$F^{(n+1)}(\xi) = 0. \tag{9}$$

En faisant $t = \xi$ dans (8) et grâce à (9) on trouve

$$F^{(n+1)}(\xi) = f^{(n+1)}(\xi) - (n+1)!\,\frac{R(x)}{\omega(x)} = 0,$$

d'où

$$R(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega(x),\quad x \in [a, b],\quad \xi \in]a, b[. \tag{10}$$

L'égalité (10) définit le reste de l'interpolation. En désignant par k la plus grande valeur de la fonction $|f^{(n+1)}(x)|$ sur $[a, b]$, on obtient la formule de la majoration du reste pour tout $x \in [a, b]$:

$$|R(x)| \leqslant \frac{k}{(n+1)!} |\omega(x)|.$$

§ 6. Méthode de calcul approché des racines d'une équation

Dans ce paragraphe on étudie le calcul approché des racines d'une équation $f(x) = 0$, où $f(x)$ est une fonction continue.

On sait comment calculer les racines de l'équation $f(x) = 0$ lorsque $f(x)$ est une fonction affine ou quadratique. Pour les fonctions plus compliquées on recourt généralement à diverses méthodes de calcul approché des racines. Etudions deux d'entre elles : la *méthode de résolution par substitution*, ou *méthode de la fourchette* et la *méthode des tangentes*, ou *méthode de Newton*.

1. Méthode de résolution par substitution. Supposons que la racine cherchée de l'équation $f(x) = 0$ est comprise dans $[a, b]$ et qu'il n'existe pas d'autres racines dans $[a, b]$. Supposons que la fonction $f(x)$ est continue sur $[a, b]$ et qu'elle prend aux bornes de cet intervalle des valeurs de signe différent. Dans la pratique, cet intervalle est déterminé par une estimation grossière. Appelons « fourchette » tout intervalle aux bornes duquel la fonction $f(x)$ prend des valeurs de signe différent.

Pour fixer les idées on admettra que $f(a) < 0$, $f(b) > 0$. Partageons $[a, b]$ en deux et considérons le segment aux bornes duquel $f(x)$ prend des valeurs de signe différent. Désignons ce segment par $[a_1, b_1]$. (Si $f(x)$ s'annule au milieu de $[a, b]$, ce milieu est la racine cherchée.) Partageons $[a_1, b_1]$ en deux et considérons celui des segments obtenus aux bornes duquel $f(x)$ a des signes différents. En poursuivant cette procédure, on obtient une suite de segments emboîtés

$$[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \ldots \supset [a_n, b_n] \ldots,$$

telle que $f(a_n) < 0$ et $f(b_n) > 0$ pour tout n. D'après le théorème 2.13 des segments emboîtés, il existe un point c intérieur à tous ces segments vers lequel convergent les suites $\{a_n\}$ et $\{b_n\}$.

Montrons que le point c est la racine cherchée, c'est-à-dire que $f(c) = 0$. La fonction $f(x)$ étant continue en c, chaque suite $\{f(a_n)\}$ et $\{f(b_n)\}$ converge vers $f(c)$. Mais des conditions $f(a_n) < 0$ et $f(b_n) > 0$ il s'ensuit, d'après le théorème 2.10, que $f(c) \leqslant 0$ et $f(c) \geqslant 0$. Donc $f(c) = 0$. C.Q.F.D.

Il est aisé de comprendre maintenant comment il faut calculer approximativement la racine $x = c$. Pour valeur approchée de

cette racine, on peut prendre le milieu du segment $[a_n, b_n]$, c'est-à-dire le point $(a_n + b_n)/2$. Vu que la longueur du segment $[a_n, b_n]$ est égale à $(b - a)/2^n$ et que la distance du point c au point $(a_n + b_n)/2$ est inférieure à la moitié de la longueur de $[a_n, b_n]$, la différence entre le nombre $(a_n + b_n)/2$ et la valeur exacte de la racine est inférieure à $(b - a)/2^{n+1}$. Cette méthode permet donc de calculer la racine cherchée c avec n'importe quelle précision si n est assez grand. Cette méthode est commode parce qu'elle implique les mêmes opérations et, par suite, est facilement programmable sur ordinateur.

2. Méthode des tangentes. Cette méthode est l'une des méthodes de calcul approché des racines de l'équation $f(x) = 0$ les plus performantes.

Supposons que la racine $x = c$ est un point intérieur au segment $[a, b]$, que la fonction $f(x)$ prend aux bornes de $[a, b]$ des valeurs de signe différent et que ses dérivées $f'(x)$ et $f''(x)$ gardent leur signe sur $[a, b]$. La dérivée $f'(x)$ ne changeant pas de signe, la fonction $f(x)$ est soit strictement croissante, soit strictement décroissante sur $[a, b]$, donc dans les deux éventualités sa courbe représentative $y = f(x)$ ne coupe l'axe Ox qu'en un seul point $x = c$ qui est l'unique racine sur $[a, b]$. De façon analogue, puisque $f''(x)$ ne change pas de signe sur $[a, b]$, la courbe $y = f(x)$ a sa convexité tournée vers le même sens sur $[a, b]$.

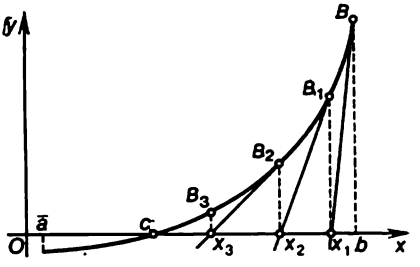


Fig. 92

Pour fixer les idées, on traitera le cas où $f'(x) > 0$ et $f''(x) > 0$. Alors $f(a) < 0$, $f(b) > 0$ et la courbe représentative est convexe vers le bas (fig. 92). Menons par le point $B(b; f(b))$ une tangente à la courbe $y = f(x)$. L'équation de cette tangente est

$$y - f(b) = f'(b)(x - b).$$

En faisant $y = 0$, on trouve l'abscisse du point d'intersection de la tangente avec l'axe Ox:

$$x_1 = b - \frac{f(b)}{f'(b)}.$$

Puisque $\frac{f(b)}{f'(b)} > 0$, on a $x_1 < b$. D'autre part, la courbe $y = f(x)$ étant située au-dessus de la tangente, on a $x_1 > c$. Donc $c < x_1 < b$. Prenons le point x_1 pour première approximation de la racine. Menons ensuite à partir du point $B_1(x_1; f(x_1))$ une tangente à la courbe $y = f(x)$ et en procédant comme plus haut pre-

nons pour deuxième approximation de la racine le point x_2:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

Ceci étant, $c < x_2 < x_1$. En poursuivant cette procédure, on obtient la formule

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \tag{1}$$

qui exprime x_{n+1} en fonction de x_n. On a donc la suite de valeurs approchées de la racine c, et de plus

$$x_1 > x_2 > x_3 > \ldots > x_n > x_{n+1} > \ldots > c. \tag{2}$$

La formule (1) est la principale formule de calcul de la méthode des tangentes. Cette méthode est une méthode d'approximations successives (d'itérations).

Montrons que la suite $\{x_n\}$ converge vers la racine cherchée c et estimons l'erreur, c'est-à-dire l'écart entre la valeur approchée x_n et la valeur exacte de la racine c. En effet, en vertu de (2) la suite $\{x_n\}$ est strictement décroissante et minorée par le nombre c. Donc d'après le théorème 2.12 elle admet une limite $c' \geqslant c$. En passant à la limite dans (1) et en tenant compte de la continuité de $f(x)$ et $f'(x)$, on obtient

$$c' = c' - \frac{f(c')}{f'(c')},$$

d'où il suit que $f(c') = 0$, c'est-à-dire que c' est une racine de l'équation $f(x) = 0$. Comme $[a, b]$ ne contient qu'une seule racine, on a $c' = c$. Donc la suite $\{x_n\}$ converge vers la racine c.

Estimons maintenant l'écart entre la n-ième approximation x_n et la racine c. En appliquant la formule de Lagrange à l'expression $f(x_n) = f(x_n) - f(c)$, on obtient $f(x_n) = (x_n - c) f'(\xi_n)$, où $c < \xi_n < x_n$. D'où la majoration

$$|x_n - c| \leqslant \frac{|f(x_n)|}{m}, \tag{3}$$

où m est la plus petite valeur de $|f'(x)|$ sur le segment $[a, b]$. La formule (3) permet d'estimer l'écart entre la valeur approchée x_n et la racine c en fonction de la valeur du module de la fonction $f(x)$ au point x_n. Signalons que la majoration (3) est valable non seulement pour la méthode des tangentes mais, d'une façon générale, pour toute méthode de calcul approché de la racine, pourvu que $m \neq 0$.

Nous avons étudié le cas où $f'(x) > 0$ et $f''(x) > 0$ sur $[a, b]$. Trois autres cas sont possibles selon les signes de $f'(x)$ et $f''(x)$: 1) $f'(x) > 0$, $f''(x) < 0$; 2) $f'(x) < 0$, $f''(x) > 0$; 3) $f'(x) < 0$, $f''(x) < 0$. Dans chacun de ces cas la méthode des tangentes se justifie comme plus haut.

Exemple. Trouver une racine de l'équation $x^2 - 5 = 0$ par la méthode des tangentes.

Solution. Considérons la fonction $f(x) = x^2 - 5$. Cette fonction est continue sur la droite numérique tout entière. Cherchons un segment aux bornes duquel $f(x)$ prend des valeurs de signe différent. Comme $f(2) = -1$ et $f(3) = 4$, ce segment est $[2, 3]$. Ce segment contient la racine cherchée. Les dérivées $f'(x) = 2x$ et $f''(x) = 2$ de $f(x)$ sont continues et strictement positives sur ce segment. Donc il faut mener la première tangente au graphique de la fonction $y = f(x)$ à partir du point $(3; 4)$. En posant $x_0 = 3$ dans la formule (1), on obtient la première approximation de la racine: $x_1 = 3 - \frac{4}{2 \cdot 3} = 2\frac{1}{3}$. En portant maintenant x_1 dans (1), on obtient $x_2 = 2\frac{1}{3} - \frac{2}{21} = 2\frac{5}{21}$ et, enfin, en portant x_2 dans (1), on obtient la troisième approximation: $x_3 = 2\frac{2}{21} - \frac{2}{987} = 2{,}23607$, et ainsi de suite.

Calculons l'erreur de l'approximation x_3 à l'aide de la formule (3). La dérivée $f'(x) = 2x$ étant strictement croissante sur $[2, 3]$, son minimum sur ce segment est $f'(2) = 4$, c'est-à-dire que $m = 4$. Calculons $f(x_3)$. On a $f(x_3) = x_3^2 - 5 = (2{,}23607)^2 - 5 = 0{,}00001$. La formule (3) nous donne maintenant

$$|x_3 - c| < \frac{0{,}00001}{4} = 0{,}0000025 = 2{,}5 \cdot 10^{-6}.$$

Si cette précision est satisfaisante, on arrête les calculs, sinon on poursuit cette procédure.

CHAPITRE VII

INTÉGRALE INDÉFINIE

§ 1. Primitive et intégrale indéfinie

1. Notion de primitive. L'un des principaux problèmes du calcul différentiel est la recherche de la dérivée d'une fonction donnée. De nombreux problèmes d'analyse mathématique et les innombrables applications de cette dernière à la géométrie, la mécanique, la physique et la technique conduisent au problème inverse, c'est-à-dire à la détermination d'une fonction $F(x)$ dont la dérivée $F'(x)$ est égale à une fonction $f(x)$ donnée.

La détermination d'une fonction dont on connaît la dérivée est un problème fondamental de calcul intégral.

Définition 1. On dit qu'une fonction $F(x)$ est une *primitive* d'une fonction $f(x)$ sur un intervalle X si $F'(x) = f(x)$ pour tout $x \in X$.

Considérons des exemples.

1. La fonction $F(x) = \sin x$ est une primitive de la fonction $f(x) = \cos x$ sur la droite numérique tout entière, puisque $(\sin x)' = \cos x$ pour tout x.

2. La fonction $F(x) = x^3$ est une primitive de la fonction $f(x) = 3x^2$ sur la droite numérique tout entière, car $(x^3)' = 3x^2$ pour tout x.

3. La fonction $F(x) = \sqrt{1 - x^2}$ est une primitive de la fonction $f(x) = -x/\sqrt{1 - x^2}$ sur l'intervalle $]-1, +1[$, puisque $(\sqrt{1 - x^2})' = -x/\sqrt{1 - x^2}$ en tout point x de cet intervalle.

La primitive d'une fonction $f(x)$ donnée n'est pas unique. En effet, si $F(x)$ est une primitive de $f(x)$, il en est de même de toute fonction $F(x) + C$, où C est une constante arbitraire, puisque $[F(x) + C]' = f(x)$. Par exemple, la fonction $f(x) = \cos x$ admet pour primitive aussi bien $\sin x$ que $\sin x + C$, puisque $(\sin x + C)' = \cos x$.

Montrons maintenant que toutes les primitives de $f(x)$ sont de la forme $F(x) + C$, où $F(x)$ est une primitive de $f(x)$.

Lemme 7.1. *Toute fonction dont la dérivée est nulle sur un intervalle X est constante sur X.*

Démonstration. Supposons que la dérivée $f'(x)$ d'une fonction $f(x)$ est nulle sur un intervalle X. D'après le théorème de Lagrange, pour deux points quelconques $x_1, x_2 \in X$, on a

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1), \quad x_1 < \xi < x_2.$$

Puisque $f'(\xi) = 0$, on a $f(x_2) = f(x_1)$. Ce qui exprime que $f(x)$ prend la même valeur sur X, autrement dit $f(x) = C$, où C est une constante. ■

Théorème 7.1. *Si $F(x)$ est une primitive d'une fonction $f(x)$ sur un intervalle X, toute autre primitive de $f(x)$ peut être mise sous la forme $F(x) + C$, où C est une constante arbitraire, sur cet intervalle.*

Démonstration. Soit $\Phi(x)$ une autre primitive de $f(x)$ sur l'intervalle X, c'est-à-dire que $\Phi'(x) = f(x)$. Alors pour tout $x \in X$

$$[\Phi(x) - F(x)]' = \Phi'(x) - F'(x) = f(x) - f(x) = 0,$$

ce qui exprime (en vertu du lemme 7.1) que la fonction $\Phi(x) - F(x) = C$, où C est une constante arbitraire. Donc $\Phi(x) = F(x) + C$. ■

De ce théorème il s'ensuit que toutes les primitives de $f(x)$ sont de la forme $F(x) + C$, où $F(x)$ est une primitive de $f(x)$ et C une constante arbitraire.

2. Intégrale indéfinie.

Définition 2. Si une fonction $F(x)$ est une primitive d'une fonction $f(x)$ sur un intervalle X, l'ensemble des fonctions $F(x) + C$, où C est une constante arbitraire, s'appelle *intégrale indéfinie de $f(x)$ sur X* et se note

$$\int f(x)\,dx = F(x) + C.$$

La fonction $f(x)$ s'appelle *intégrant*, la variable x, *variable d'intégration*.

Le symbole $\int f(x)\,dx$ désigne donc l'ensemble de toutes les primitives de la fonction $f(x)$. On comprendra parfois par ce symbole un élément quelconque de cet ensemble, c'est-à-dire une primitive.

La détermination d'une fonction par sa dérivée s'appelle *intégration* de cette fonction. L'intégration est l'opération inverse de la dérivation. Pour vérifier qu'une intégration est correcte, il suffit de dériver le résultat obtenu pour obtenir l'intégrant.

Dans ce chapitre on glisse sur le problème de l'existence des primitives (donc des intégrales indéfinies) pour de vastes classes de fonctions. Signalons qu'au chapitre VIII, § 6, on prouvera que

toute fonction continue sur un segment possède une primitive (donc une intégrale indéfinie) sur ce segment.

Exemples.

1. $\int 3x^2\,dx = x^3 + C$, puisque $(x^3 + C)' = 3x^2$.

2. $\int \cos x\,dx = \sin x + C$, puisque $(\sin x + C)' = \cos x$.

3. $\int \frac{1}{x}\,dx = \ln|x| + C$, puisque $(\ln|x| + C)' = \frac{1}{x}$.

4. $\int e^{-2x}\,dx = -\frac{1}{2}e^{-2x} + C$, puisque $\left(-\frac{1}{2}e^{-2x} + C\right)' = e^{-2x}$, etc.

§ 2. Propriétés fondamentales de l'intégrale indéfinie

Les propriétés suivantes de l'intégrale indéfinie sont une conséquence immédiate de sa définition.

1° La dérivée d'une intégrale indéfinie est égale à l'intégrant ; la différentielle d'une intégrale indéfinie, à l'expression sous le signe somme, c'est-à-dire

$$\left(\int f(x)\,dx\right)' = f(x) \quad \text{et} \quad d\int f(x)\,dx = f(x)\,dx.$$

En effet, $\left(\int f(x)\,dx\right)' = (F(x) + C)' = F'(x) = f(x)$ et

$$d\int f(x)\,dx = \left(\int f(x)\,dx\right)' dx = f(x)\,dx.$$

2° L'intégrale indéfinie de la différentielle d'une fonction est égale à la somme de cette fonction et d'une constante arbitraire, c'est-à-dire

$$\int dF(x) = F(x) + C.$$

En effet, puisque $dF(x) = F'(x)\,dx$, il vient $\int F'(x)\,dx = F(x) + C$.

3° Tout facteur constant peut être sorti du signe somme, c'est-à-dire si $k = \text{const}$, alors

$$\int kf(x)\,dx = k\int f(x)\,dx.$$

En effet, soit $F(x)$ une primitive de $f(x)$: $F'(x) = f(x)$. Alors $kF(x)$ est une primitive de $kf(x)$: $(kF(x))' = kF'(x) = kf(x)$. D'où

$$k\int f(x)\,dx = k[F(x) + C] = kF(x) + C_1 = \int kf(x)dx,$$

où $C_1 = kC$.

4° L'intégrale indéfinie de la somme algébrique de deux fonctions est égale à la somme algébrique des intégrales indéfinies de

ces fonctions, c'est-à-dire

$$\int f(x) \pm g(x)\,dx = \int f(x)\,dx \pm \int g(x)\,dx.$$

En effet, soient $F(x)$ et $G(x)$ des primitives de $f(x)$ et $g(x)$:

$$F'(x) = f(x),\ G'(x) = g(x).$$

Les fonctions $F(x) \pm G(x)$ sont alors des primitives des fonctions $f(x) \pm g(x)$. Donc

$$\int f(x)\,dx \pm \int g(x)\,dx = [F(x) + C_1] \pm [G(x) + C_2] =$$
$$= [F(x) \pm G(x)] + [C_1 \pm C_2] = [F(x) \pm G(x)] + C =$$
$$= \int [f(x) \pm g(x)]\,dx.$$

Cette propriété est valable pour tout nombre fini de fonctions.

§ 3. Table des principales intégrales

Les principales intégrales sont regroupées dans la table suivante. Une partie des formules résultent directement de la définition de l'intégration comme l'opération inverse de la dérivation et de la table des dérivées. La véracité des autres formules se vérifie aisément par dérivation.

I. $\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C$ $(\alpha \neq -1)$.

II. $\int \frac{dx}{x} = \ln|x| + C$.

III. $\int \frac{dx}{1+x^2} = \operatorname{Arc\,tg} x + C$.

IV. $\int \frac{dx}{\sqrt{1-x^2}} = \operatorname{Arc\,sin} x + C$.

V. $\int a^x\,dx = \frac{a^x}{\ln a} + C$ $(0 < a \neq 1)$.

VI. $\int e^x\,dx = e^x + C$.

VII. $\int \sin x\,dx = -\cos x + C$.

VIII. $\int \cos x\,dx = \sin x + C$.

IX. $\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C$.

X. $\int \frac{dx}{\sin^2 x} = -\operatorname{cotg} x + C$.

XI. $\int \frac{dx}{x^2-a^2} = \frac{1}{2a} \ln\left|\frac{x-a}{x+a}\right| + C$ $(a \neq 0)$.

XII. $\int \frac{dx}{\sqrt{x^2+k}} = \ln|x + \sqrt{x^2+k}| + C$.

XIII. $\int \frac{dx}{x^2+a^2} = \frac{1}{a} \times \operatorname{Arc\,tg} \frac{x}{a} + C$.

XIV. $\int \frac{dx}{\sqrt{a^2-x^2}} = \operatorname{Arc\,sin} \frac{x}{a} + C$.

Les intégr ales de cette table s'appellent *tabulaires.*

§ 4. Principales méthodes d'intégration

1. Intégration directe. Le calcul des intégrales par une utilisation directe de la table des intégrales élémentaires et des propriétés fondamentales des intégrales indéfinies s'appelle *intégration directe.*

Exemple 1. $\int \left(5\cos x + 2 - 3x^2 + \frac{1}{x} - \frac{4}{x^2+1}\right) dx =$

$= 5\int \cos x\,dx + 2\int dx - 3\int x^2\,dx + \int \frac{dx}{x} - 4\int \frac{dx}{x^2+1} =$

$= 5\sin x + 2x - x^3 + \ln|x| - 4\operatorname{Arc\,tg} x + C.$

Exemple 2. $\int \left(\sin\frac{x}{2} + \cos\frac{x}{2}\right)^2 dx = \int \left(\sin^2\frac{x}{2} + 2\sin\frac{x}{2} \times\right.$

$\left. \times \cos\frac{x}{2} + \cos^2\frac{x}{2}\right) dx = \int (1+\sin x)\,dx = \int dx + \int \sin x\,dx =$

$= x - \cos x + C.$

Exemple 3. $\int \operatorname{tg}^2 x\,dx = \int (\sec^2 x - 1)\,dx = \int \frac{dx}{\cos^2 x} - \int dx =$

$= \operatorname{tg} x - x + C.$

2. Méthode d'intégration par substitution. Dans nombre de cas l'introduction d'une nouvelle variable d'intégration permet de ramener l'intégration à la recherche d'une intégrale tabulaire, c'est-à-dire de se ramener à une intégration directe. Cette méthode qui s'appelle *méthode d'intégration par substitution*, ou encore *méthode de changement de la variable*, est basée sur le théorème suivant.

Théorème 7.2. *Soit $x = \varphi(t)$ une fonction définie et dérivable sur un intervalle T et soit X son ensemble des valeurs. Soit enfin $f(x)$ une fonction définie sur X. Si la fonction $f(x)$ admet une primitive sur X, alors*

$$\int f(x)\,dx\Big|_{x=\varphi(t)} = \int f[\varphi(t)]\,\varphi'(t)\,dt \qquad (1)$$

sur T.

Démonstration. Soit $F(x)$ une primitive de $f(x)$ sur X. Considérons la fonction composée $F[\varphi(t)]$ sur T. La règle de dérivation d'une fonction composée nous donne

$$(F[\varphi(t)])' = F'_x[\varphi(t)]\,\varphi'(t) = f[\varphi(t)]\,\varphi'(t),$$

c'est-à-dire que $F[\varphi(t)]$ est une primitive de la fonction $f[\varphi(t)]\varphi'(t)$ sur T, donc

$$\int f[\varphi(t)]\varphi'(t)\,dt = F[\varphi(t)] + C.$$

En remarquant que $F[\varphi(t)] + C = (F(x) + C)|_{x=\varphi(t)} = \int f(x)\,dx\Big|_{x=\varphi(t)}$, on obtient la formule (1). ■

La formule (1) s'appelle *formule de changement de la variable d'intégration.*

Exemple 1. Calculer l'intégrale $\int \frac{x^3}{(x-1)^2}\,dx$.

Solution. Posons $x - 1 = t$; alors $x = t + 1$. D'où $dx = dt$. La formule (1) nous donne

$$\int \frac{x^3}{(x-1)^2}\,dx = \int \frac{(t+1)^3}{t^2}\,dt = \int \left(t + 3 + \frac{3}{t} + \frac{1}{t^2}\right) dt =$$
$$= \frac{1}{2}t^2 + 3t + 3\ln|t| - \frac{1}{t} + C.$$

En revenant à la variable x, on obtient en définitive

$$\int \frac{x^3}{(x-1)^2}\,dx = \frac{1}{2}(x-1)^2 + 3(x-1) + 3\ln|x-1| - \frac{1}{x-1} + C.$$

Remarque. Quand on change la variable d'intégration dans une intégrale indéfinie, on a parfois intérêt à se donner non pas x en fonction de t mais t en fonction de x.

Exemple 2. Calculer l'intégrale $\int \frac{x^4\,dx}{x^5+7}$.

Solution. Posons $x^5 + 7 = t$, $dt = 5x^4\,dx$; alors

$$\int \frac{x^4\,dx}{x^5+7} = \frac{1}{5}\int \frac{dt}{t} = \frac{1}{5}\ln|t| + C,$$

si bien que

$$\int \frac{x^4\,dx}{x^5+7} = \frac{1}{5}\ln|x^5+7| + C.$$

Signalons que le choix convenable de la substitution soulève de sérieuses difficultés. Pour les surmonter il faut être doué d'une bonne technique de dérivation et bien connaître les intégrales tabulaires.

Exemple 3. Calculer l'intégrale $\int \frac{dx}{\sqrt{x^2+a}}$.

Solution. Posons $\sqrt{x^2+a} + x = t$, d'où $\left(\frac{x}{\sqrt{x^2+a}} + 1\right) \times dx = dt$. Donc

$$dx = \frac{\sqrt{x^2+a}}{\sqrt{x^2+a}+x}\,dt,$$

de sorte que

$$\int \frac{dx}{\sqrt{x^2+a}} = \int \frac{dt}{t} = \ln|t| + C = \ln|\sqrt{x^2+a} + x| + C.$$

Exemple 4. Calculer l'intégrale $\int \sin^n x \cos x \, dx$.

Solution. Posons $t = \sin x$, $dt = \cos x \, dx$. Alors

$$\int \sin^n x \cos x \, dx = \int t^n \, dt = \begin{cases} \frac{t^{n+1}}{n+1} + C = \frac{\sin^{n+1} x}{n+1} + C \\ \qquad\qquad \text{pour } n \neq -1, \\ \ln|t| + C = \ln|\sin x| + C \\ \qquad\qquad \text{pour } n = -1. \end{cases}$$

Exemple 5. Calculer l'intégrale $\int \frac{x \, dx}{(x^2+1)^n}$, $n \neq 1$.

Solution. Posons $x^2 + 1 = t$, $2x \, dx = dt$; alors

$$\int \frac{x \, dx}{(x^2+1)^n} = \frac{1}{2} \int \frac{dt}{t^n} = -\frac{1}{2(n-1)} \frac{1}{t^{n-1}} + C =$$

$$= -\frac{1}{2(n-1)} \frac{1}{(x^2+1)^{n-1}} + C.$$

Pour $n = 1$ on obtient par analogie

$$\int \frac{x \, dx}{x^2+1} = \frac{1}{2} \ln(x^2+1) + C.$$

3. Méthode d'intégration par parties. La méthode d'intégration par parties est basée sur la formule de dérivation du produit de deux fonctions.

Théorème 7.3. *Soient $u(x)$ et $v(x)$ des fonctions définies et dérivables sur un intervalle X et supposons que la fonction $u'(x)\, v(x)$ admet une primitive sur X. Alors la fonction $u(x)\, v'(x)$ admet aussi une primitive sur X et*

$$\int u(x)\, v'(x)\, dx = u(x)\, v(x) - \int v(x)\, u'(x)\, dx. \qquad (2)$$

Démonstration. La relation

$$[u(x)\, v(x)]' = u'(x)\, v(x) + u(x)\, v'(x)$$

entraîne

$$u(x)\, v'(x) = [u(x)\, v(x)]' - u'(x)\, v(x).$$

La fonction $u(x)\, v(x)$ est une primitive de la fonction $[u(x)\, v(x)]'$ sur X. La fonction $u'(x)\, v(x)$ admet une primitive sur X par

hypothèse. Donc la fonction $u(x)\,v'(x)$ en admet une aussi sur X. Une intégration de la dernière égalité nous donne la formule (2). ■

La formule (2) s'appelle *formule d'intégration par parties.*

Comme $v'(x)\,dx = dv$, $u'(x)\,dx = du$, on peut la mettre sous la forme

$$\int u\,dv = u\cdot v - \int v\,du.$$

Cette formule permet de ramener le calcul de $\int u\,dv$ à celui de $\int v\,du$ qui est parfois plus simple.

Exemple 1.

$$\int \underbrace{\operatorname{Arc\,tg} x}_{u}\,\underbrace{dx}_{} = \left| \begin{array}{l} u = \operatorname{Arc\,tg} x,\ du = \dfrac{dx}{1+x^2} \\ dv = dx,\ \int dv = \int dx,\ v = x \end{array} \right| =$$

$$= \underbrace{x}_{v}\,\underbrace{\operatorname{Arc\,tg} x}_{u} - \int \underbrace{x}_{v}\,\underbrace{\frac{dx}{1+x^2}}_{du} = x\operatorname{Arc\,tg} x -$$

$$-\frac{1}{2}\int \frac{d(x^2+1)}{1+x^2} = x\operatorname{Arc\,tg} x - \frac{1}{2}\ln(1+x^2) + C.$$

Exemple 2.

$$\int xe^x\,dx = \left| \begin{array}{l} u = x,\ du = dx \\ dv = e^x\,dx,\ \int dv = \int e^x\,dx,\ v = e^x \end{array} \right| =$$

$$= \int \underbrace{x}_{u}\,\underbrace{de^x}_{dv} = \underbrace{x}_{u}\,\underbrace{e^x}_{v} - \int \underbrace{e^x}_{v}\,\underbrace{dx}_{du} = xe^x - e^x + C.$$

Exemple 3.

$$\int x\ln x\,dx = \left| \begin{array}{l} u = \ln x,\ du = \dfrac{dx}{x} \\ dv = x\,dx,\ \int dv = \int x\,dx,\ v = \dfrac{x^2}{2} \end{array} \right| =$$

$$= \int \ln x\,d\left(\frac{x^2}{2}\right) = \frac{x^2}{2}\ln x - \int \frac{x^2}{2}\,\frac{1}{x}\,dx = \frac{x^2}{2}\ln x - \frac{x^2}{4} + C.$$

On est amené parfois à se servir plusieurs fois de la formule d'intégration par parties pour calculer une intégrale.

Exemple 4.

$$\int x^2 \cos x\,dx =$$

$$= \left| \begin{array}{l} u = x^2, \ du = 2x\,dx \\ dv = \cos x\,dx, \ \int dv = \int \cos x\,dx, \quad v = \sin x \end{array} \right| =$$

$$= \int x^2\,d(\sin x) = x^2 \sin x - 2\int x \sin x\,dx =$$

$$= \left| \begin{array}{l} u = x, \ du = dx \\ dv = \sin x\,dx, \ \int dv = \int \sin x\,dx, \quad v = -\cos x \end{array} \right| =$$

$$= x^2 \sin x + 2\int x\,d\cos x = x^2 \sin x + 2x\cos x - 2\int \cos x\,dx =$$

$$= x^2 \sin x + 2x\cos x - 2\sin x + C.$$

L'intégrale $\int x^2 \cos x\,dx$ a été donc calculée par une double intégration par parties.

Calculons en conclusion l'intégrale

$$I_n = \int \frac{dx}{(x^2+1)^n}$$

(n est un entier >0) qui nous sera utile dans le paragraphe suivant. Pour $n = 1$, on a l'intégrale tabulaire

$$I_1 = \text{Arc tg}\, x + C.$$

Supposons que $n > 1$. En représentant 1 au dénominateur comme la différence $(x^2 + 1) - x^2$, on obtient

$$I_n = \int \frac{dx}{(x^2+1)^{n-1}} - \int \frac{x^2\,dx}{(x^2+1)^n}.$$

Calculons la deuxième intégrale par parties :

$$u = x, \ du = dx, \ dv = \frac{x\,dx}{(x^2+1)^n}, \ v = \int \frac{x\,dx}{(x^2+1)^n} =$$

$$= -\frac{1}{(2n-2)(x^2+1)^{n-1}}$$

(cf. n° 2, exemple 5), alors

$$\int \frac{x^2\,dx}{(x^2+1)^n} = -\frac{x}{(2n-2)(x^2+1)^{n-1}} + \int \frac{dx}{(2n-2)(x^2+1)^{n-1}},$$

donc

$$I_n = I_{n-1} + \frac{x}{(2n-2)(x^2+1)^{n-1}} - \frac{1}{2n-2} I_{n-1},$$

d'où

$$I_n = \frac{x}{(2n-2)(x^2+1)^{n-1}} + \frac{2n-3}{2n-2} I_{n-1}.$$

L'intégrale I_n est donc exprimée en fonction de I_{n-1} :

$$\int \frac{dx}{(x^2+1)^n} = \frac{x}{(2n-2)(x^2+1)^{n-1}} + \frac{2n-3}{2n-2} \int \frac{dx}{(x^2+1)^{n-1}}, \quad n > 1. \quad (3)$$

Les formules de type (3) s'appellent *formules de récurrence.*

Exemple 5. Calculer $\int \frac{dx}{(x^2+1)^3}$.

Solution. La formule de récurrence (3) nous donne

$$\int \frac{dx}{(x^2+1)^3} = \frac{x}{4(x^2+1)^2} + \frac{3}{4} \int \frac{dx}{(x^2+1)^2},$$

$$\int \frac{dx}{(x^2+1)^2} = \frac{x}{2(x^2+1)} + \frac{1}{2} \int \frac{dx}{x^2+1},$$

or

$$I_1 = \int \frac{dx}{x^2+1} = \operatorname{Arc\,tg} x,$$

donc

$$\int \frac{dx}{(x^2+1)^3} = \frac{x}{4(x^2+1)^2} + \frac{3x}{8(x^2+1)} + \frac{3}{8} \operatorname{Arc\,tg} x + C.$$

§ 5. Intégration des fonctions rationnelles

Les fractions (ou fonctions) rationnelles, c'est-à-dire les fonctions de la forme

$$\frac{P(x)}{Q(x)}$$

où $P(x)$ et $Q(x)$ sont des fonctions polynômes, forment une importante classe de fonctions dont les intégrales s'expriment toujours par l'intermédiaire de fonctions élémentaires.

Si le degré du numérateur est supérieur au degré du dénominateur, on obtient alors

$$\frac{P(x)}{Q(x)} = W(x) + \frac{R(x)}{Q(x)}, \quad (1)$$

où $W(x)$ est un polynôme appelé *partie entière* de la fraction $\frac{P(x)}{Q(x)}$ et $\frac{R(x)}{Q(x)}$, une fraction irréductible.

Exemples.

1. $\frac{x^5+x^3-x^2+1}{x^3-2x+1} = x^2+3-\frac{2x^2-6x+2}{x^3-2x+1}$.

2. $\frac{x^3+x+1}{x^2+1} = x + \frac{1}{x^2+1}$.

En algèbre supérieure on démontre que tout polynôme peut être décomposé en un produit de la forme

$$Q(x) = A(x-\alpha)(x-\beta)\ldots(x-\gamma),$$

où A est le coefficient dominant de $Q(x)$, $\alpha, \beta, \ldots, \gamma$ les racines de l'équation $Q(x) = 0$. Les facteurs $(x - \alpha), (x - \beta), \ldots, (x - \gamma)$ s'appellent *éléments simples*. Si des éléments simples sont confondus, on obtient la décomposition

$$Q(x) = A(x - \alpha)^r (x - \beta)^s \ldots (x - \gamma)^t, \qquad (2)$$

où $r, s, \ldots, t$ sont des entiers appelés *multiplicités des racines* $\alpha, \beta, \ldots, \gamma$ respectivement, et tels que $r + s + \ldots + t = n$, où n est le degré de $Q(x)$.

Le polynôme $Q(x)$ peut présenter aussi des racines complexes. On démontre en algèbre que si $\alpha = a + bi$ est une racine complexe de multiplicité r d'un polynôme à coefficients réels, il en est de même de sa conjuguée complexe $\bar{\alpha} = a - bi$. En d'autres termes, si la décomposition (2) contient le facteur $(x - \alpha)^r$, où $\alpha = a + bi$ $(b \neq 0)$, elle contiendra également le facteur $(x - \bar{\alpha})^r$. En multipliant ces deux facteurs, on obtient

$$\begin{aligned}(x - \alpha)^r (x - \bar{\alpha})^r &= \{[x - (a + bi)][x - (a - bi)]\}^r = \\ &= [x^2 - x(a + bi) - x(a - bi) + a^2 + b^2]^r = \\ &= [x^2 - 2ax + a^2 + b^2]^r = (x^2 + 2px + q)^r,\end{aligned}$$

où $p = -a$, $q = a^2 + b^2$, $p^2 - q < 0$, p et q sont réels.

En procédant de même avec les autres racines complexes, on peut mettre la décomposition (2) sous la forme

$$\begin{aligned}Q(x) = A(x - \alpha)^r (x - \beta)^s \ldots (x^2 + 2px + q)^t (x^2 + \\ + 2ux + v)^\eta \ldots, \qquad (3)\end{aligned}$$

où $\alpha, \beta, \ldots, p, q, u, v, \ldots$ sont des réels.

Le théorème suivant est prouvé en algèbre supérieure.

Théorème. *Une fraction rationnelle* $\frac{R(x)}{Q(x)}$ *irréductible dont le dénominateur* $Q(x)$ *est un polynôme de la forme* (3), *se représente d'une manière unique sous la forme*

$$\begin{aligned}\frac{R(x)}{Q(x)} = &\frac{A_1}{(x-\alpha)} + \frac{A_2}{(x-\alpha)^2} + \ldots + \frac{A_r}{(x-\alpha)^r} + \ldots \\ &\ldots + \frac{M_1 x + N_1}{x^2 + 2px + q} + \frac{M_2 x + N_2}{(x^2 + 2px + q)^2} + \ldots \\ &\ldots + \frac{M_t x + N_t}{(x^2 + 2px + q)^t} + \ldots, \qquad (4)\end{aligned}$$

où $A_1, A_2, \ldots, A_r, \ldots, M_1, N_1, M_2, N_2, \ldots, M_t, N_t, \ldots$ *sont des réels.*

L'expression (4) s'appelle *décomposition d'une fraction rationnelle en éléments simples.*

La décomposition (4) a lieu pour tout x qui n'est pas un zéro réel de $Q(x)$.

Pour déterminer les nombres $A_1, A_2, \ldots, A_r, \ldots$ $\ldots, M_1, N_1, \ldots, M_t, N_t, \ldots$, multiplions les deux parties de (4) par $Q(x)$. Puisque l'égalité entre le polynôme $R(x)$ et le polynôme obtenu au second membre doit avoir lieu pour tout x, les coefficients de même rang doivent être égaux. Leur identification nous conduit à un système d'équations du premier degré dont la résolution nous donne les inconnues $A_1, A_2, \ldots, A_r, \ldots$ $\ldots, M_1, N_1, \ldots, M_t, N_t, \ldots$

Cette méthode de détermination des coefficients s'appelle *méthode des coefficients indéterminés.*

Exemple 1. Décomposer la fraction rationnelle $\frac{2x-1}{x^2-5x+6}$ en éléments simples.

Solution. Puisque $x^2 - 5x + 6 = (x - 3)(x - 2)$, la formule (4) nous donne

$$\frac{2x-1}{x^2-5x+6} = \frac{A}{x-3} + \frac{B}{x-2}.$$

En multipliant les deux membres de cette égalité par $x^2 - 5x + 6$, on obtient

$$2x - 1 = A(x - 2) + B(x - 3) \text{ ou } 2x - 1 = (A + B)x - 2A - 3B.$$

L'identification des coefficients de même rang nous conduit au système d'équations du premier degré en A et B:

$$\begin{cases} A + B = 2, \\ 2A + 3B = 1, \end{cases}$$

d'où $A = 5$, $B = -3$.

Donc

$$\frac{2x-1}{x^2-5x+6} = \frac{5}{x-3} - \frac{3}{x-2}.$$

Exemple 2. Décomposer la fraction rationnelle $\frac{x^2-1}{x(x^2+1)^2}$ en éléments simples.

Solution. Les racines du binôme $x^2 + 1$ étant complexes, la formule (4) entraîne

$$\frac{x^2-1}{x(x^2+1)^2} = \frac{A}{x} + \frac{Bx+C}{x^2+1} + \frac{Dx+E}{(x^2+1)^2}.$$

Une multiplication des deux membres de cette égalité par $x(x^2 + 1)^2$ nous donne

$$x^2 - 1 = A(x^2 + 1)^2 + (Bx + C)(x^2 + 1)x + (Dx + E)x$$

ou

$$x^2 - 1 = (A + B)\,x^4 + Cx^3 + (2A + B + D)\,x^2 + \\ + (C + E)\,x + A.$$

Une identification des coefficients de même rang nous conduit au système d'équations:

$$\begin{cases} x^4 : A + B = 0, \\ x^3 : C = 0, \\ x^2 : 2A + B + D = 0, \\ x^1 : C + E = 0, \\ x^0 : A = -1, \end{cases}$$

dont la résolution nous donne $A = -1$, $B = 1$, $C = 0$, $D = 2$, $E = 0$, donc la décomposition cherchée est

$$\frac{x^2-1}{x(x^2+1)^2} = -\frac{1}{x} + \frac{x}{x^2+1} + \frac{2x}{(x^2+1)^2}.$$

De ce qui précède il s'ensuit que l'intégration d'une fonction rationnelle (1) se ramène à celle d'un polynôme $W(x) = a_0x^m + a_1x^{m-1} + \ldots + a_m$ dont l'intégrale est tabulaire:

$$\int W(x)\,dx = a_0\frac{x^{m+1}}{m+1} + a_1\frac{x^m}{m} + \cdots + a_mx + C,$$

et à celle d'une fraction rationnelle irréductible $\frac{R(x)}{Q(x)}$; ceci nous conduit au calcul d'intégrales des quatre types suivants *):

I. $\int \frac{A}{x-\alpha}\,dx = A \ln|x-\alpha| + C$;

II. $\int \frac{A}{(x-\alpha)^r}\,dx = -\frac{A}{(r-1)(x-\alpha)^{r-1}} + C \quad (r > 1)$;

III. $\int \frac{Ax+B}{x^2+2px+q}\,dx$;

IV. $\int \frac{Ax+B}{(x^2+2px+q)^r}\,dx \quad (r > 1)$.

Ceci étant, le trinôme $x^2 + 2px + q$ n'admet pas de racines réelles, c'est-à-dire que $p^2 - q < 0$.

Calculons les intégrales III que l'on rencontre fréquemment en pratique. Complétons le trinôme $x^2 + 2px + q$ à un carré:

$$x^2 + 2px + q = (x + p)^2 + q - p^2.$$

Cette représentation nous suggère la substitution $x + p = t$, d'où $x = t - p$, $dx = dt$. Posons par ailleurs $q - p^2 = h > 0$ et pas-

*) Les intégrales I et II se calculent de façon élémentaire par la substitution $t = x - \alpha$.

sons à la variable t. L'intégrale se ramène à la forme

$$\int \frac{Ax+B}{x^2+2px+q}\,dx = \int \frac{At+B-Ap}{t^2+h}\,dt = \\ = \frac{1}{2}A\int \frac{2t\,dt}{t^2+h} + (B-A\cdot p)\int \frac{dt}{t^2+h}.$$

Le calcul de la première intégrale de droite est immédiat:

$$\int \frac{2t\,dt}{t^2+h} = \ln|t^2+h| + C = \ln|x^2+2px+q| + C.$$

La deuxième intégrale est tabulaire.

Exemple 3. Calculer $\int \frac{6x+5}{x^2+4x+9}\,dx$.

Solution. Complétons le dénominateur à un carré: $x^2 + 4x + 9 = (x+2)^2 + 5$. Faisons la substitution $x + 2 = t$. On en déduit $x = t - 2$, $dx = dt$, donc

$$\int \frac{6x+5}{x^2+4x+9}\,dx = \int \frac{6x+5}{(x+2)^2+5}\,dx = \int \frac{6(t-2)+5}{t^2+5}\,dt = \\ = \int \frac{6t-7}{t^2+5}\,dt = 3\int \frac{2t\,dt}{t^2+5} - 7\int \frac{dt}{t^2+5} = \\ = 3\ln(t^2+5) - \frac{7}{\sqrt{5}}\operatorname{Arc\,tg}\frac{t}{\sqrt{5}} + C.$$

En revenant à la variable x, on obtient

$$\int \frac{6x+5}{x^2+4x+9}\,dx = 3\ln(x^2+4x+9) - \frac{7}{\sqrt{5}}\operatorname{Arc\,tg}\frac{x+2}{\sqrt{5}} + C.$$

Calculons maintenant une intégrale IV: $\int \frac{Ax+B}{(x^2+2px+q)^r}\,dx$, $q-p^2>0$, $r>1$. Introduisons à cet effet la nouvelle variable z par la formule

$$z = (x+p)/\sqrt{q-p^2},\ \text{d'où}\ x = z\sqrt{q-p^2} - p,\ dx = \\ = \sqrt{q-p^2}\,dz. \quad (5)$$

On a par ailleurs

$$z^2+1 = \frac{(x+p)^2}{q-p^2} + 1 = \frac{x^2+2px+q}{q-p^2}. \quad (6)$$

En faisant la substitution (5) et en tenant compte de (6), on obtient donc

$$\int \frac{Ax+B}{(x^2+2px+q)^r}\,dx = \int \frac{A[z\sqrt{q-p^2}-p]+B}{(z^2+1)^r(q-p^2)^r}\sqrt{q-p^2}\,dz = \\ = \int \frac{Mz+N}{(z^2+1)^r}\,dz = M\int \frac{z\,dz}{(z^2+1)^r} + N\int \frac{dz}{(z^2+1)^r},$$

où M et N se déduisent de l'avant-dernière égalité.

La deuxième intégrale peut être calculée à l'aide de la formule de récurrence [cf. § 4, n° 3, formule (3)]. En posant $z^2 + 1 = t$ dans la première intégrale, on obtient

$$M \int \frac{z\,dz}{(z^2+1)^r} = \frac{M}{2} \int \frac{dt}{t^r} = -\frac{M}{2(r-1)} \frac{1}{t^{r-1}} + C =$$

$$= -\frac{M}{2(r-1)} \frac{1}{(z^2+1)^{r-1}} + C.$$

Exemple 4. Calculer $\int \frac{5x+3}{(x^2-2x+5)^2}\,dx$.

Solution. Posons $z = \frac{x-1}{\sqrt{5-1}} = \frac{x-1}{2}$, d'où $x = 1 + 2z$, $dx = 2\,dz$ et $x^2 - 2x + 5 = 4(z^2+1)$, donc

$$\int \frac{5x+3}{(x^2-2x+5)^2}\,dx = \int \frac{5(1+2z)+3}{4^2(z^2+1)^2}\,2dz = \int \frac{10z+8}{8(z^2+1)^2}\,dz =$$

$$= \frac{5}{4} \int \frac{z\,dz}{(z^2+1)^2} + \int \frac{dz}{(z^2+1)^2}.$$

Mais

$$\int \frac{z\,dz}{(z^2+1)^2} = -\frac{1}{2}\frac{1}{z^2+1}, \quad \int \frac{dz}{(z^2+1)^2} = \frac{z}{2(z^2+1)} + \frac{1}{2} \operatorname{Arc\,tg} z.$$

Donc

$$\int \frac{5x+3}{(x^2-2x+5)^2}\,dx = -\frac{5}{8}\frac{1}{z^2+1} + \frac{z}{2(z^2+1)} + \frac{1}{2} \operatorname{Arc\,tg} z + C =$$

$$= \frac{4z-5}{8(z^2+1)} + \frac{1}{2} \operatorname{Arc\,tg} z + C.$$

En revenant à la variable x, on obtient

$$\int \frac{5x+3}{(x^2-2x+5)^2}\,dx = \frac{2x-7}{2(x^2-2x+5)} + \frac{1}{2} \operatorname{Arc\,tg}\left(\frac{x-1}{2}\right) + C.$$

On a ainsi établi que l'intégration de toute fraction rationnelle se ramène à celle d'un polynôme et d'un nombre fini d'éléments simples dont les intégrales s'expriment par des fonctions rationnelles, logarithmes et des arcs tangentes. Autrement dit, toute fraction rationnelle s'intègre par des fonctions élémentaires.

§ 6. Intégration des fonctions irrationnelles et transcendantes

Désignons préalablement par $R(u, v)$ une fonction rationnelle de deux variables u et v, c'est-à-dire une fonction obtenue uniquement par des opérations arithmétiques (addition, soustraction, multiplication et division) sur des variables u et v. Exemple :

$$R(u, v) = \frac{3u^2v + u^5v^4}{u^3 + 4v^2}.$$

Si les variables u et v sont à leur tour des fonctions de x: $u = \varphi(x)$, $v = \psi(x)$, la fonction $R[\varphi(x), \psi(x)]$ s'appelle fonction rationnelle de $\varphi(x)$ et $\psi(x)$. Par exemple, la fonction

$$f(x) = \frac{x + \sqrt{(x^2-1)^2}}{x^2 - 5\sqrt{x^2-1}}$$

est une fonction rationnelle de x et de $\sqrt{x^2-1}$: $f(x) = R(x, \sqrt{x^2-1})$. Ici $R(u, v) = \frac{u+v^2}{u^2-5v}$, $u = x$, $v = \sqrt{x^2-1}$. La fonction

$$f(x) = \frac{\sin^2 x - \cos^2 x}{\sin^3 x + 2\cos x}$$

est une fonction rationnelle de $\sin x$ et $\cos x$: $f(x) = R(\sin x, \cos x)$.

Considérons maintenant les intégrales de certaines fonctions irrationnelles et transcendantes élémentaires et montrons qu'elles se ramènent à des intégrales de fonctions rationnelles qui peuvent être calculées par les méthodes du § 5.

1. Intégrale de la forme $\int R\left(x, \sqrt[m]{\frac{ax+b}{cx+d}}\right) dx$, où a, b, c et d sont des nombres arbitraires $\left(\frac{a}{c} \neq \frac{b}{d}\right)$; m, un entier naturel, R une fonction rationnelle de x et $\sqrt[m]{\frac{ax+b}{cx+d}}$. Montrons que la substitution

$$t = \sqrt[m]{\frac{ax+b}{cx+d}}$$

nous ramène à l'intégration d'une fonction rationnelle. En effet,

$$t^m = \frac{ax+b}{cx+d}, \quad x = \frac{b-dt^m}{ct^m-a}, \quad dx = \frac{mt^{m-1}(ad-bc)}{(ct^m-a)^2}\,dt,$$

de sorte que

$$\int R\left(x, \sqrt[m]{\frac{ax+b}{cx+d}}\right) dx = \int R\left(\frac{b-dt^m}{ct^m-a}, t\right) \frac{mt^{m-1}(ad-bc)'}{(ct^m-a)^2}\,dt = \int R_1(t)\,dt,$$

où $R_1(t)$ est une fonction rationnelle de t.

Exemple 1. Calculer $\int \sqrt{\frac{1+x}{1-x}} \frac{dx}{1-x}$.

Solution. En faisant la substitution $t = \sqrt{\frac{1+x}{1-x}}$, on obtient $t^2 = \frac{1+x}{1-x}$, $1-x = \frac{2}{t^2+1}$, $x = \frac{t^2-1}{t^2+1}$, $dx = \frac{4t\,dt}{(t^2+1)^2}$.

On a ensuite

$$\int \sqrt{\frac{1+x}{1-x}}\,\frac{dx}{1-x} = 2\int \frac{t^2\,dt}{t^2+1} = 2\int \frac{t^2+1-1}{t^2+1}\,dt = 2\int dt - 2\int \frac{dt}{t^2+1} = 2t - 2\operatorname{Arc\,tg} t + C = 2\sqrt{\frac{1+x}{1-x}} - 2\operatorname{Arc\,tg}\sqrt{\frac{1+x}{1-x}} + C.$$

Exemple 2. Calculer $\int \frac{dx}{\sqrt{x}+\sqrt[3]{x}}$.

Solution. On a

$$\int \frac{dx}{\sqrt{x}+\sqrt[3]{x}} = \int \frac{dx}{(\sqrt[6]{x})^2+(\sqrt[6]{x})^3} = \left|\begin{matrix} t=\sqrt[6]{x},\ x=t^6 \\ dx=6t^5\,dt \end{matrix}\right| = \int \frac{6t^5\,dt}{t^2+t^3} = 6\int \frac{t^3\,dt}{t+1} = 6\int\left[(t^2-t+1)\,dt - \int \frac{dt}{t+1}\right] = 6\left[\frac{t^3}{3} - \frac{t^2}{2} + t - \ln|t+1|\right] + C = 2\sqrt{x} - 3\sqrt[3]{x} + 6\sqrt[6]{x} - 6\ln(\sqrt[6]{x}+1) + C.$$

2. Intégrale de la forme $\int R(x, \sqrt{ax^2+bx+c})\,dx$, où a, b et c sont des nombres arbitraires; $a \neq 0$; R une fonction rationnelle de x et $\sqrt{ax^2+bx+c}$.

Si le trinôme ax^2+bx+c possède des racines réelles x_1 et x_2 distinctes et $a > 0$, alors

$$\sqrt{ax^2+bx+c} = \sqrt{a(x-x_1)(x-x_2)} = |x-x_1|\sqrt{\frac{a(x-x_2)}{x-x_1}}.$$

Donc

$$R(x, \sqrt{ax^2+bx+c}) = R\left(x, |x-x_1|\sqrt{\frac{a(x-x_2)}{x-x_1}}\right) = R_1\left(x, \sqrt{\frac{a(x-x_2)}{x-x_1}}\right),$$

c'est-à-dire qu'on obtient l'intégrale du n° 1.

Si $x_1 = x_2$, alors

$$\sqrt{ax^2+bx+c} = |x-x_1|\sqrt{a},$$

c'est-à-dire qu'on a une fonction rationnelle de x sous le signe d'intégration.

Il est donc plus intéressant de traiter le cas où le trinôme ax^2+bx+c n'admet pas de racines réelles et $a > 0$. Montrons que dans ce cas on se ramène à l'intégration d'une fonction ration-

nelle par la substitution d'Euler *)

$$t = \sqrt{ax^2+bx+c} + x\sqrt{a}. \text{ **)}$$

En élevant les deux membres de l'égalité $\sqrt{ax^2+bx+c} = t - x\sqrt{a}$ au carré, on obtient $bx+c = t^2 - 2\sqrt{a}\,tx$, de sorte que

$$x = \frac{t^2-c}{2\sqrt{a}t+b}, \quad \sqrt{ax^2+bx+c} = \frac{\sqrt{a}t^2+bt+c\sqrt{a}}{2\sqrt{a}t+b},$$

$$dx = 2\,\frac{\sqrt{a}t^2+bt+c\sqrt{a}}{(2\sqrt{a}t+b)^2}\,dt.$$

Donc

$$\int R\left(x, \sqrt{ax^2+bx+c}\right)dx =$$

$$= \int R\left(\frac{t^2-c}{2\sqrt{a}t+b}, \frac{\sqrt{a}t^2+bt+c\sqrt{a}}{2\sqrt{a}t+b}\right) 2\,\frac{\sqrt{a}t^2+bt+c\sqrt{a}}{(2\sqrt{a}t+b)^2}\,dt =$$

$$= \int R_1(t)\,dt,$$

où $R_1(t)$ est une fonction rationnelle de t.

Si $a < 0$ et $c > 0$, on utilise une autre substitution d'Euler

$$\sqrt{ax^2+bx+c} = xt \pm \sqrt{c}.$$

Exemple 3. Calculer $\int \frac{dx}{x+\sqrt{x^2+x+1}}$.

Solution. Les racines du trinôme $x^2 + x + 1$ étant complexes, faisons la substitution $\sqrt{x^2+x+1} = t - x$. En élevant au carré, on obtient

$$x^2 + x + 1 = t^2 - 2tx + x^2$$

ou

$$x + 1 = t^2 - 2tx,$$

d'où

$$x = \frac{t^2-1}{1+2t}, \quad dx = 2\,\frac{t^2+t+1}{(1+2t)^2}\,dt.$$

Alors

$$\int \frac{dx}{x+\sqrt{x^2+x+1}} = 2\int \frac{t^2+t+1}{t(1+2t)^2}\,dt.$$

Par ailleurs,

$$\frac{2t^2+2t+2}{t(1+2t)^2} = \frac{A}{t} + \frac{B}{1+2t} + \frac{D}{(1+2t)^2}.$$

*) Euler Leonhard (1707-1783), mathématicien, mécanicien, physicien et astronome suisse.

**) On peut poser $t = \sqrt{ax^2+bx+c} - x\sqrt{a}$.

Une multiplication des deux membres de cette égalité par $t(1+2t)^2$ nous donne

$$2t^2+2t+2=A(1+2t)^2+Bt(1+2t)+Dt,$$

ou

$$2t^2+2t+2=(4A+2B)t^2+(4A+D+B)t+A.$$

Une identification des coefficients de t nous conduit au système d'équations du premier degré en A, B et D :

$$\begin{cases} 4A+2B=2, \\ 4A+D+B=2, \\ A=2 \end{cases}$$

d'où $A=2$, $B=-3$, $D=-3$. Donc

$$\frac{2t^2+2t+2}{t(1+2t)^2}=\frac{2}{t}-\frac{3}{1+2t}-\frac{3}{(1+2t)^2},$$

et en définitive

$$\int\frac{dx}{x+\sqrt{x^2+x+1}}=\int\left[\frac{2}{t}-\frac{3}{1+2t}-\frac{3}{(1+2t)^2}\right]dt=$$

$$=2\int\frac{dt}{t}-\frac{3}{2}\int\frac{d(1+2t)}{1+2t}-\frac{3}{2}\int\frac{d(1+2t)}{(1+2t)^2}=$$

$$=2\ln|t|-\frac{3}{2}\ln|1+2t|+\frac{3}{2(1+2t)}+C=$$

$$=2\ln|\sqrt{x^2+x+1}+x|-\frac{3}{2}\ln|1+2x+$$

$$+2\sqrt{x^2+x+1}|+\frac{3}{2(1+2x+2\sqrt{x^2+x+1})}+C.$$

Exemple 4. Calculer $\int\frac{dx}{(1+x)\sqrt{1+x-x^2}}$.

Solution. Les racines de $1+x-x^2$ étant complexes et $a<0$, $c>0$, on utilisera la substitution $\sqrt{1+x-x^2}=tx-1$. En élevant les deux membres au carré, on obtient

$$1+x-x^2=t^2x^2-2tx+1 \text{ ou } 1-x=t^2x-2t,$$

d'où

$$x=\frac{1+2t}{t^2+1},\quad dx=\frac{2(1-t-t^2)}{(t^2+1)^2}dt,\quad \sqrt{1+x-x^2}=\frac{t^2+t-1}{t^2+1}.$$

Donc

$$\int\frac{dx}{(1+x)\sqrt{1+x-x^2}}=\int\frac{2(1-t-t^2)}{\left(1+\frac{1+2t}{t^2+1}\right)\frac{t^2+t-1}{t^2+1}(t^2+1)^2}dt=$$

$$=-2\int\frac{dt}{1+(t+1)^2}=-2\operatorname{Arc\,tg}(t+1)+C=$$

$$=-2\operatorname{Arc\,tg}\frac{\sqrt{1+x-x^2}+x+1}{x}+C.$$

Signalons que le calcul des intégrales par la substitution d'Euler conduit généralement à des calculs laborieux et des expressions volumineuses, aussi n'y recourt-on que dans les cas où il est impossible de calculer ces intégrales par un procédé plus court.

3. Intégrale de la forme $\int R(\sin x, \cos x)\,dx$, où R est une fonction rationnelle de $\sin x$ et $\cos x$. Montrons que la substitution

$$t = \operatorname{tg}\frac{x}{2}, \quad x \in [-\pi, \pi]$$

nous ramène à l'intégration d'une fonction rationnelle. En effet,

$$\sin x = \frac{2\operatorname{tg}(x/2)}{1+\operatorname{tg}^2(x/2)} = \frac{2t}{1+t^2}, \quad \cos x = \frac{1-\operatorname{tg}^2(x/2)}{1+\operatorname{tg}^2(x/2)} = \frac{1-t^2}{1+t^2};$$

$$x = 2\operatorname{Arc}\operatorname{tg} t, \quad dx = \frac{2\,dt}{1+t^2},$$

de sorte que

$$\int R(\sin x, \cos x)\,dx = \int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)\frac{2\,dt}{1+t^2} = \int R_1(t)\,dt,$$

où $R_1(t)$ est une fonction rationnelle de t.

Exemple 5. Calculer $\int \frac{dx}{1+\sin x}$.

Solution. La substitution $t = \operatorname{tg}\frac{x}{2}$ nous donne

$$\sin x = \frac{2t}{1+t^2}, \quad x = 2\operatorname{Arc}\operatorname{tg} t, \quad dx = \frac{2\,dt}{1+t^2}.$$

Donc

$$\int \frac{dx}{1+\sin x} = 2\int \frac{dt}{(1+t)^2} = -\frac{2}{1+t} + C = -\frac{2}{1+\operatorname{tg}(x/2)} + C.$$

4. Intégrale de la forme $\int R(e^x)\,dx$. Montrons que la substitution

$$t = e^x$$

nous conduit à l'intégration d'une fonction rationnelle. En effet, vu que $x = \ln t$ et $dx = \frac{dt}{t}$, on obtient

$$\int R(e^x)\,dx = \int R(t)\frac{dt}{t},$$

où $R(t)$ est une fonction rationnelle de t.

Exemple 6. Calculer $\int \frac{e^x-1}{e^x+1}\,dx$.

Solution. Posons $t=e^x$. D'où $dx=\frac{dt}{t}$. Donc

$$\int \frac{e^x-1}{e^x+1}\,dx=\int \frac{t-1}{t+1}\frac{dt}{t}=\int \frac{2t-(t+1)}{(t+1)t}\,dt=2\int \frac{dt}{t+1}-\int\frac{dt}{t}=$$
$$=2\ln(1+t)-\ln t+C=2\ln(1+e^x)-x+C.$$

Signalons en conclusion que les méthodes d'intégration envisagées n'épuisent pas toutes les classes de fonctions élémentaires analytiquement intégrables. On a pu se rendre compte que l'intégration est plus compliquée que la dérivation. Il faut une certaine technique et une certaine ingéniosité qui s'acquièrent par la résolution d'un grand nombre d'exercices.

Signalons par ailleurs que si la dérivation est une opération stable, il en va autrement de l'intégration. Il existe des fonctions élémentaires (par exemple e^{-x^2}, $\frac{1}{\ln x}$, $\frac{\sin x}{x}$, etc.) dont les primitives ne sont pas des fonctions élémentaires. Non seulement ces primitives existent, mais elles jouent un rôle important tant en analyse mathématique que dans ses applications. Ces primitives sont bien étudiées et leurs tables et graphiques permettent de les utiliser fréquemment en pratique.

CHAPITRE VIII

INTÉGRALE DÉFINIE

§ 1. Définition de l'intégrale définie

Soit $y = f(x)$ une fonction définie sur un intervalle $[a, b]$, $a < b$. Subdivisons cet intervalle en n parties arbitraires à l'aide des points :

$$a = x_0 < x_1 < x_2 < \ldots < x_{i-1} < x_i < \ldots < x_n = b.$$

Désignons cette subdivision par τ et appelons les points $x_0, x_1, \ldots, x_n$ points de la subdivision τ. Prenons un point arbitraire ξ_i dans chaque intervalle partiel $[x_{i-1}, x_i]$. Désignons par Δx_i la longueur $x_i - x_{i-1}$ de l'intervalle $[x_{i-1}, x_i]$. Formons la somme

$$\sigma = f(\xi_1)\,\Delta x_1 + f(\xi_2)\,\Delta x_2 + \ldots + f(\xi_n)\,\Delta x_n = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i \qquad (1)$$

et appelons-la *somme intégrale* de $f(x)$ attachée à la subdivision τ et au choix des points ξ_i. La signification géométrique de la somme σ est évidente : c'est l'aire des rectangles de bases $\Delta x_1, \Delta x_2, \ldots, \Delta x_n$ et de hauteurs $f(\xi_1), f(\xi_2), \ldots, f(\xi_n)$ (fig. 93) (si $f(x) \geqslant 0$).

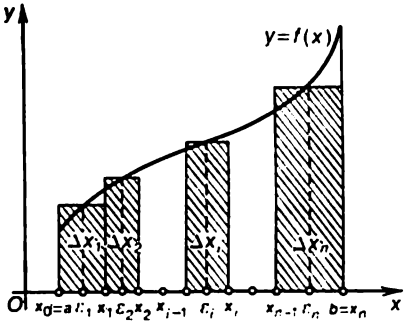


Fig. 93

Appelons *pas* de la subdivision τ et désignons par λ le plus grand nombre Δx_i : $\lambda = \max_{1 \leqslant i \leqslant n} \{\Delta x_i\}$.

Définition. La limite finie I (si elle existe) de la somme intégrale (1) lorsque $\lambda \to 0$ s'appelle *intégrale définie de la fonction* $f(x)$ *sur l'intervalle* $[a, b]$ et se note :

$$I = \int_a^b f(x)\,dx \qquad (2)$$

ou

$$\int_a^b f(x)\,dx = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i.$$

La fonction $f(x)$ est dite *intégrable* sur $[a, b]$. Les nombres a et b s'appellent respectivement *limites* (ou *bornes*) *inférieure* et *supérieure* d'intégration, la fonction $f(x)$, *intégrant*, la variable x, *variable d'intégration*.

Faisons quelques éclaircissements, puisque nous avons affaire à un passage à la limite qui n'est pas ordinaire. En effet, la somme intégrale dépend des points de subdivision x_i et des points intermédiaires ξ_i. Le nombre des points x_i et ξ_i tend vers l'infini lorsque $\lambda \to 0$. Donc la notion de limite d'une somme intégrale appelle quelques précisions. Définissons d'abord cette notion dans le langage des suites. Soient $\{\tau_k\}$ des subdivisions d'un intervalle $[a, b]$ telles que le pas λ_k de τ_k tend vers 0 lorsque $k \to \infty$. Choisissons dans chaque subdivision des points intermédiaires arbitraires ξ_i. On obtient ainsi une suite de subdivisions $\{\tau_k\}$ telle que $\lim\limits_{k \to \infty} \lambda_k = 0$.

On dit qu'une fonction $f(x)$ est *intégrable sur* $[a, b]$ si pour toute suite de subdivisions $\{\tau_k\}$ telle que $\lim\limits_{k \to \infty} \lambda_k = 0$ la suite $\{\sigma_k\}$ des sommes intégrales correspondantes tend vers un même nombre I.

On peut donner une définition de l'intégrale définie en termes de ε, δ : on dit qu'un nombre I est l'*intégrale définie d'une fonction* $f(x)$ *sur un intervalle* $[a, b]$ si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour $\lambda < \delta$ l'on ait

$$\left| \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i - I \right| < \varepsilon$$

indépendamment du choix des points ξ_i.

La démonstration de l'équivalence de ces deux définitions est calquée sur celle de l'équivalence des définitions de la limite d'une fonction. La définition en termes de suites permet d'étendre les principales notions de la théorie des limites à cette nouvelle forme de limite.

De la définition de l'intégrale définie il s'ensuit que la valeur de l'intégrale (2) dépend uniquement de la forme des fonctions $f(x)$ et des nombres a et b. Si donc la fonction $f(x)$ et les bornes d'intégration sont données, l'intégrale (2) est définie de façon unique. Il s'ensuit en particulier qu'une intégrale définie ne dépend pas du choix de l'argument de l'intégrant, c'est-à-dire du choix de la variable d'intégration :

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt = \int_a^b f(\xi)\,d\xi, \text{ etc.}$$

§ 2. Conditions d'existence de l'intégrale définie

1. Limitation de l'intégrant.

Théorème 8.1 (condition nécessaire d'intégrabilité d'une fonction). *Si une fonction $f(x)$ est intégrable sur un intervalle $[a, b]$, elle est bornée sur cet intervalle.*

Démonstration. Supposons par absurde que la fonction $f(x)$ n'est pas bornée sur $[a, b]$. Montrons que pour toute subdivision de l'intervalle $[a, b]$ la somme intégrale σ peut être rendue arbitrairement grande par le seul choix des points $\xi_1, \xi_2, \ldots, \xi_n$.

En effet, la fonction $f(x)$ étant infinie sur $[a, b]$, elle l'est au moins sur un intervalle partiel, $[x_0, x_1]$ pour fixer les idées, quelle que soit la subdivision envisagée. Choisissons des points arbitraires $\xi_2, \xi_3, \ldots, \xi_n$ dans les autres intervalles partiels et posons

$$\sigma' = f(\xi_2)\,\Delta x_2 + f(\xi_3)\,\Delta x_3 + \ldots + f(\xi_n)\,\Delta x_n.$$

Considérons un nombre arbitraire $M > 0$ et prenons un $\xi_1 \in [x_0, x_1]$ tel que

$$|f(\xi_1)| \geqslant \frac{|\sigma'| + M}{\Delta x_1}.$$

Ceci est possible puisque $f(x)$ est infinie sur $[x_0, x_1]$. Alors

$$|f(\xi_1)|\,\Delta x_1 \geqslant |\sigma'| + M \text{ et } |\sigma| = |f(\xi_1)\,\Delta x_1 + \sigma'| \geqslant$$
$$\geqslant |f(\xi_1)|\,\Delta x_1 - |\sigma'| \geqslant M,$$

c'est-à-dire que la somme intégrale σ est supérieure en valeur absolue à tout nombre donné à l'avance. Donc la somme intégrale σ ne possède pas de limite finie lorsque $\lambda \to 0$, or ceci exprime que l'intégrale définie d'une fonction infinie n'existe pas. ■

Remarque. La réciproque n'est pas vraie, c'est-à-dire que la condition de limitation de la fonction $f(x)$ est une condition nécessaire mais pas suffisante d'intégrabilité. Montrons-le sur un exemple. Considérons la fonction de Dirichlet sur l'intervalle $[0, 1]$:

$$f(x) = \begin{cases} 1 & \text{si } x \text{ est rationnel} \\ 0 & \text{si } x \text{ est irrationnel.} \end{cases}$$

La fonction de Dirichlet est visiblement bornée. Mais elle n'est pas intégrable sur $[0, 1]$. Prouvons-le. Soit une subdivision quelconque de l'intervalle $[0, 1]$. Si les points intermédiaires ξ_i sont choisis rationnels, on a

$$\sigma = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = \sum_{i=1}^{n} 1 \cdot \Delta x_i = 1,$$

s'ils sont irrationnels,

$$\sigma = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = \sum_{i=1}^{n} 0 \cdot \Delta x_i = 0.$$

Donc la somme intégrale σ peut prendre aussi bien la valeur 0 que la valeur 1, autrement dit elle n'admet pas de limite lorsque $\lambda \to 0$.

Donc la limitation d'une fonction $f(x)$ ne suffit pas à elle seule à assurer l'intégrabilité de cette fonction. Il faut d'autres propriétés dont l'établissement implique l'introduction des notions de sommes inférieures et supérieures de Darboux.

2. Sommes de Darboux *). Soient $f(x)$ une fonction bornée sur un intervalle $[a, b]$, τ une subdivision de $[a, b]$: $a = x_0 < x_1 < \ldots < x_{i-1} < x_i < \ldots < x_n = b$. Désignons par m_i et M_i l'infimum et le suprémum respectifs de cette fonction sur $[x_{i-1}, x_i]$ et formons les sommes suivantes:

$$S = M_1\Delta x_1 + M_2\Delta x_2 + \ldots + M_n\Delta x_n = \sum_{i=1}^{n} M_i\Delta x_i,$$

$$s = m_1\Delta x_1 + m_2\Delta x_2 + \ldots + m_n\Delta x_n = \sum_{i=1}^{n} m_i\Delta x_i.$$

Ces sommes s'appellent respectivement *sommes supérieure* et *inférieure (de Darboux)* de la fonction $f(x)$ attachées à la subdivision τ de l'intervalle $[a, b]$.

De la définition de l'infimum et du suprémum, il s'ensuit que $m_i \leqslant f(\xi_i) \leqslant M_i$ pour $\xi_i \in [x_{i-1}, x_i]$. D'où

$$s = \sum_{i=1}^{n} m_i\Delta x_i \leqslant \sigma = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i \leqslant \sum_{i=1}^{n} M_i\Delta x_i = S,$$

c'est-à-dire que la somme intégrale et les sommes de Darboux attachées à une subdivision donnée sont reliées par la double inégalité

$$s \leqslant \sigma \leqslant S. \tag{1}$$

Les sommes de Darboux admettent une interprétation géométrique simple. Considérons une fonction $f(x)$ continue positive sur $[a, b]$ et le trapèze curviligne correspondant, c'est-à-dire la figure limitée par le graphique de la fonction $f(x)$, les droites verticales $x = a$, $x = b$ et l'axe Ox (fig. 94 et 95). La fonction $f(x)$ est continue sur $[x_{i-1}, x_i]$, car elle l'est sur $[a, b]$. D'après le deuxième théorème de Weierstrass elle atteint son infimum et son suprémum sur $[x_{i-1}, x_i]$ et par suite m_i et M_i sont respectivement

*) Darboux Gaston (1842-1917), mathématicien français.

le minimum et le maximum de $f(x)$ sur cet intervalle. Donc les sommes S et s sont égales aux aires des figures hachurées respectivement des figures 94 et 95.

Signalons tout particulièrement que les sommes de Darboux dépendent seulement de la subdivision de l'intervalle $[a, b]$, tandis que la somme intégrale σ dépend encore du choix des points intermédiaires ξ_i des intervalles partiels $[x_{i-1}, x_i]$. Si l'on

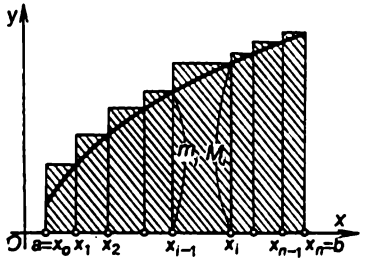


Fig. 94

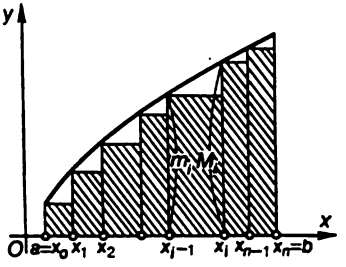


Fig. 95

fixe une subdivision de $[a, b]$, les sommes s et S sont des nombres définis, alors que la somme σ est une quantité variable, puisque les points ξ_i sont arbitraires.

3. Propriétés des sommes de Darboux. 1° *Pour toute subdivision fixe* τ *et tout* $\varepsilon > 0$, *les points* $\xi_i \in [x_{i-1}, x_i]$ *peuvent être choisis de telle sorte que la somme intégrale* σ *vérifie la double inégalité* $0 \leqslant S - \sigma < \varepsilon$. *Les points* ξ_i *peuvent aussi être choisis de telle sorte que la somme intégrale* σ *satisfasse la double inégalité* $0 \leqslant \sigma - s < \varepsilon$.

Démonstration. Soit τ une subdivision fixe de $[a, b]$. Prouvons par exemple la double inégalité $0 \leqslant S - \sigma < \varepsilon$. D'après la propriété du suprémum M_i, pour tout $\varepsilon > 0$ donné on peut exhiber un point ξ_i de $[x_{i-1}, x_i]$ tel que

$$0 \leqslant M_i - f(\xi_i) < \frac{\varepsilon}{b-a}, \quad i = 1, 2, \ldots, n.$$

En multipliant ces doubles inégalités par Δx_i et en ajoutant on obtient $0 \leqslant S - \sigma < \varepsilon$. On établit de façon analogue la double inégalité $0 \leqslant \sigma - s < \varepsilon$. ■

2° *L'adjonction de nouveaux points à une subdivision* τ *donnée de l'intervalle* $[a, b]$ *ne diminue pas la somme inférieure de Darboux et n'augmente pas la somme supérieure.*

Démonstration. On peut se borner à l'adjonction d'un seul point à la subdivision τ, car si on en a plusieurs, on peut les ajouter un par un. Ajoutons un point x' à τ et supposons qu'il tombe dans l'intervalle $[x_{i-1}, x_i]$ (fig. 96). Désignons respective-

ment par s et s' les sommes inférieures de Darboux et par S et S', les sommes supérieures attachées à τ et à la nouvelle subdivision τ'.

Faisons la démonstration pour les sommes inférieures s et s'. Désignons par m_i' et m_i'' les infimums de la fonction $f(x)$ respectivement sur les intervalles $[x_{i-1}, x']$ et $[x', x_i]$. Les sommes s et s' contiennent les mêmes termes à l'exception de deux: $m_i \Delta x_i$

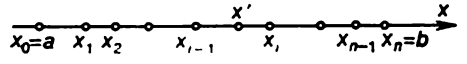


Fig. 96

et $m_i'(x' - x_{i-1}) + m_i''(x_i - x')$. Comme $m_i' \geqslant m_i$ et $m_i'' \geqslant m_i$ (l'infimum sur une partie de $[x_{i-1}, x_i]$ est $\geqslant$ à l'infimum sur l'intervalle $[x_{i-1}, x_i]$ tout entier), il vient

$$m_i'(x' - x_{i-1}) + m_i''(x_i - x') \geqslant m_i(x' - x_{i-1}) + m_i(x_i - x') = m_i \Delta x_i.$$

D'où $s' \geqslant s$.

On démontre de façon analogue que $S' \leqslant S$. ■

3°. *La somme inférieure de Darboux attachée à toute subdivision τ' est inférieure à la somme supérieure attachée à toute autre subdivision τ''.*

Démonstration. Soient s', S' et s'', S'' les sommes inférieures et supérieures de Darboux attachées respectivement aux subdivisions τ' et τ''. Considérons la subdivision τ composée de tous les points de τ' et τ''. Désignons par s et S les sommes de Darboux attachées à τ. La subdivision τ ayant été obtenue par l'adjonction des points de τ'' à τ', on a en vertu de la propriété 2° et compte tenu de l'inégalité évidente $s \leqslant S$:

$$s' \leqslant s \leqslant S \leqslant S'.$$

Mais on peut aussi obtenir τ en ajoutant à τ'' les points de τ'. Donc

$$s'' \leqslant s \leqslant S \leqslant S''.$$

En comparant ces inégalités, on trouve que $s' \leqslant S''$, $s'' \leqslant S'$. ■

4° *Les ensembles $\{S\}$ et $\{s\}$ respectivement des sommes supérieures et inférieures de Darboux d'une fonction $f(x)$ attachées à toutes les subdivisions d'un intervalle $[a, b]$ sont respectivement minoré et majoré et de plus la borne supérieure de $\{s\}$ est inférieure à la borne inférieure de $\{S\}$.*

Démonstration. Cette propriété résulte directement de la propriété 3°. En effet, l'ensemble $\{S\}$ de toutes les sommes supérieures de Darboux est minoré, par exemple par toute somme

inférieure s de Darboux, tandis que l'ensemble $\{s\}$ de toutes les sommes inférieures de Darboux est majoré, par exemple par toute somme supérieure S de Darboux. Donc les ensembles $\{S\}$ et $\{s\}$ admettent des bornes d'après le théorème 1.1. Désignons par I^* la borne inférieure de $\{S\}$ et par I_* la borne supérieure de $\{s\}$:

$$I^* = \inf \{S\}, \quad I_* = \sup \{s\}.$$

Montrons que $I_* \leqslant I^*$. Supposons par absurde que $I_* > I^*$. Posons $I_* - I^* = \varepsilon > 0$. Des propriétés des bornes I_* et I^* il s'ensuit qu'il existe des nombres S' et s'', respectivement sommes supérieure et inférieure de Darboux attachées à des subdivisions τ' et τ'' de $[a, b]$, tels que $I^* + \varepsilon/2 > S'$ et $I_* - \varepsilon/2 < s''$. En soustrayant la deuxième inégalité de la première on trouve que $S' - s'' < I^* - I_* + \varepsilon$, or $I^* - I_* = -\varepsilon$, donc $S' - s'' < 0$, c'est-à-dire que $s'' > S'$. Cette contradiction avec la propriété 3° montre que $I_* \leqslant I^*$. ■

4. Condition nécessaire et suffisante d'intégrabilité. On a le théorème fondamental suivant.

Théorème 8.2. *Pour qu'une fonction $f(x)$ bornée sur un intervalle $[a, b]$ soit intégrable sur cet intervalle, il est nécessaire et suffisant que*

$$\lim_{\lambda \to 0} (S - s) = 0. \tag{2}$$

La condition (2) exprime que pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que $| S - s | < \varepsilon$ pour $\lambda < \delta$. Puisque $s \leqslant S$, cette inégalité est équivalente à l'inégalité

$$S - s < \varepsilon. \tag{3}$$

Démonstration. *Nécessité.* Supposons que la fonction $f(x)$ est intégrable sur $[a, b]$, c'est-à-dire qu'existe l'intégrale définie $I = \int_a^b f(x)\,dx$. Ceci exprime que pour tout $\varepsilon > 0$ il existe un nombre $\delta > 0$ tel que pour toute subdivision τ dont le pas $\lambda < \delta$ l'on a

$$| \sigma - I | < \varepsilon/4 \tag{4}$$

indépendamment du choix des points ξ_i. Considérons une telle subdivision τ. D'après la propriété 1°, on peut exhiber des sommes intégrales σ' et σ'' attachées à τ telles que

$$S - \sigma' \leqslant \varepsilon/4, \quad \sigma'' - s \leqslant \varepsilon/4. \tag{5}$$

A noter que les sommes intégrales σ' et σ'' vérifient toutes deux l'inégalité (4). La relation

$$S - s = (S - \sigma') + (\sigma' - I) + (I - \sigma'') + (\sigma'' - s)$$

et les inégalités (4) et (5) entraînent

$$S - s < \varepsilon.$$

Suffisance. Soit réalisée la condition (3). D'après la propriété 4°, pour toutes sommes inférieures et supérieures de Darboux on a $s \leqslant I_* \leqslant I^* \leqslant S$, donc $0 \leqslant I^* - I_* \leqslant S - s$, d'où il s'ensuit, en vertu de (3), que $0 \leqslant I^* - I_* < \varepsilon$ pour tout $\varepsilon > 0$. Ce qui exprime que $I^* - I_* = 0$, c'est-à-dire que $I^* = I_*$. En posant $I = I^* = I_*$ on trouve que

$$s \leqslant I \leqslant S \tag{6}$$

pour toute subdivision. Si une somme intégrale σ et les sommes de Darboux s et S sont attachées à une même subdivision τ, alors [cf. formule (1)]

$$s \leqslant \sigma \leqslant S. \tag{7}$$

Les doubles inégalités (6) et (7) entraînent

$$|\sigma - I| \leqslant S - s. \tag{8}$$

Par hypothèse, pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que l'inégalité (3) est vérifiée pour $\lambda < \delta$. Il s'ensuit alors de l'égalité (8) que

$$|\sigma - I| < \varepsilon \text{ pour } \lambda < \delta,$$

ce qui exprime que le nombre I est la limite de la somme intégrale σ lorsque $\lambda \to 0$, c'est-à-dire que la fonction $f(x)$ est intégrable sur $[a, b]$. ■

Nous aurons besoin dans la suite d'une autre forme d'écriture de la condition nécessaire et suffisante d'intégrabilité. En désignant par ω_i l'oscillation $M_i - m_i$ de la fonction $f(x)$ sur l'intervalle $[x_{i-1}, x_i]$, on obtient

$$S - s = \sum_{i=1}^{n} M_i \Delta x_i - \sum_{i=1}^{n} m_i \Delta x_i = \sum_{i=1}^{n} (M_i - m_i)\, \Delta x_i = \sum_{i=1}^{n} \omega_i \Delta x_i.$$

Chaque terme de la dernière somme est positif, puisque $M_i \geqslant m_i$ et $\Delta x_i > 0$ et la condition d'existence de l'intégrale définie peut s'énoncer comme suit: pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que

$$\sum_{i=1}^{n} \omega_i \Delta x_i < \varepsilon \quad \text{pour} \quad \lambda < \delta.$$

Cette condition est généralement utilisée sous cette forme.

§ 3. Intégrabilité des fonctions continues et de certaines fonctions discontinues

Théorème 8.3. *Si une fonction $f(x)$ est continue sur un intervalle $[a, b]$, elle est intégrable sur cet intervalle.*

Démonstration. Etant continue sur $[a, b]$, la fonction $f(x)$ y est uniformément continue d'après le théorème de Cantor. Soit donné $\varepsilon > 0$. D'après le corollaire du théorème de Cantor, pour tout nombre $\varepsilon/(b-a)$ il existe un $\delta > 0$ tel que les oscillations ω_i soient $< \varepsilon/(b-a)$ pour toute subdivision de $[a, b]$ en intervalles partiels $[x_{i-1}, x_i]$ de longueur $\Delta x_i < \delta$. D'où

$$S - s = \sum_{i=1}^{n} \omega_i \Delta x_i < \frac{\varepsilon}{b-a} \sum_{i=1}^{n} \Delta x_i = \varepsilon \quad \text{pour} \quad \lambda < \delta.$$

La condition suffisante d'intégrabilité est donc remplie pour la fonction $f(x)$ continue sur $[a, b]$. Cette condition entraîne l'existence de l'intégrale définie. ■

Le théorème nous dit que la condition de continuité d'une fonction est une condition suffisante d'intégrabilité de cette fonction. Mais cela ne signifie pas que l'intégrale définie n'existe

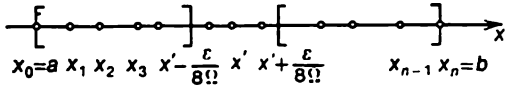


Fig. 97

que pour les fonctions continues. La classe des fonctions intégrables est bien plus large. Cette classe contient par exemple des fonctions présentant un nombre fini de discontinuités. Prouvons-le.

Théorème 8.4. *Si une fonction $f(x)$ est bornée et continue partout sauf en un nombre fini de points d'un intervalle $[a, b]$, alors elle est intégrable sur cet intervalle.*

Démonstration. Il suffit de traiter le cas où l'intervalle $[a, b]$ ne contient qu'un seul point de discontinuité x'. Soient M et m les bornes de $f(x)$ sur $[a, b]$, $\Omega = M - m$ l'oscillation de $f(x)$ sur cet intervalle. Prenons un $\varepsilon > 0$ arbitraire assez petit et considérons les intervalles $\left[a, x' - \frac{\varepsilon}{8\Omega}\right]$ et $\left[x' + \frac{\varepsilon}{8\Omega}, b\right]$ (fig. 97). La fonction $f(x)$ étant continue sur chacun de ces intervalles, il existe un $\delta' > 0$ tel que pour toute subdivision de ces intervalles en intervalles partiels $[x'_{i-1}, x'_i]$ de longueur $\Delta x'_i < \delta'$, les oscillations ω_i sont $< \frac{\varepsilon}{2(b-a)}$.

Soit $\delta = \min\left\{\delta', \frac{\varepsilon}{8\Omega}\right\}$. Considérons maintenant une subdivision quelconque de $[a, b]$ en intervalles partiels de longueur $\Delta x_i < \delta$ (fig. 97). Décomposons la somme $\sum_{i=1}^{n} \omega_i \Delta x_i$ attachée à cette subdivision en la somme $\sum' \omega_i \Delta x_i + \sum'' \omega_i \Delta x_i$: le premier terme contient des intervalles partiels strictement extérieurs à l'$\frac{\varepsilon}{8\Omega}$-voisinage du point x', le second, des intervalles partiels qui soit sont strictement intérieurs à l'$\frac{\varepsilon}{8\Omega}$-voisinage du point x', soit possèdent des points communs avec ce voisinage.

Pour la première somme, comme dans la démonstration du théorème précédent, on a

$$\sum' \omega_i \Delta x_i < \varepsilon/2;$$

s'agissant de la deuxième somme, on remarquera que la somme des longueurs des intervalles strictement intérieurs à l'$\frac{\varepsilon}{8\Omega}$-voisinage de x' est $\leqslant \frac{\varepsilon}{4\Omega}$; deux intervalles partiels au plus sont partiellement contenus dans ce voisinage, donc la somme de leurs longueurs est $< 2\delta \leqslant \frac{\varepsilon}{4\Omega}$. Donc

$$\sum'' \omega_i \Delta x_i \leqslant \Omega \sum'' \Delta x_i < \Omega \frac{\varepsilon}{2\Omega} = \frac{\varepsilon}{2}.$$

En définitive,

$$S - s = \sum_{i=1}^{n} \omega_i \Delta x_i =$$

$$= \sum' \omega_i \Delta x_i + \sum'' \omega_i \Delta x_i < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \text{pour} \quad \lambda < \delta.$$

Ce qui prouve l'intégrabilité de $f(x)$ sur $[a, b]$. ■

Corollaire. *Une fonction continue par morceaux sur un segment est intégrable sur ce segment.*

§ 4. Propriétés fondamentales de l'intégrale définie

1° L'intégrale $\int_a^b f(x)\,dx$ a été introduite dans le cas où $a < b$. Généralisons la notion d'intégrale définie au cas où les bornes d'intégration sont confondues ou bien la borne inférieure est

supérieure à la borne supérieure. *On admet par définition que*

$$\int_a^a f(x)\,dx = 0, \tag{1}$$

cette formule étant considérée comme la généralisation naturelle de l'intégrale définie à un segment de longueur nulle.

De même, *on admet par définition que*

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx, \tag{2}$$

cette formule étant considérée comme la généralisation naturelle de la notion d'intégrale définie au cas où l'intervalle $[a, b]$, $a < b$, est parcouru de b vers a. Dans ce cas les points de subdivision x_i de $[a, b]$ sont numérotés de b vers a et les différences $\Delta x_i = x_i - x_{i-1}$ de la somme intégrale sont toutes négatives.

2° *Quels que soient les nombres a, b et c, on a*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx. \tag{3}$$

(Dans les §§ 4, 5 on admet l'existence des intégrales intervenant dans les formules démontrées.)

Démonstration. Supposons d'abord que $a < c < b$. La limite de la somme intégrale σ étant indépendante de la subdivision de $[a, b]$, on choisira cette substitution de telle sorte que c en soit un point. Si par exemple $c = x_m$, on peut décomposer la somme intégrale σ en deux sommes:

$$\sigma = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = \sum_{i=1}^{m} f(\xi_i)\,\Delta x_i + \sum_{i=m+1}^{n} f(\xi_i)\,\Delta x_i.$$

En faisant tendre λ vers 0, on obtient l'égalité (3).

La propriété prouvée exprime que l'intégrale définie sur un intervalle est égale à la somme des intégrales sur les parties de cet intervalle.

Pour toute autre disposition des points a, b et c la démonstration se ramène facilement au cas étudié. Supposons par exemple que $a < b < c$. D'après ce qui a été prouvé, on a alors

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx,$$

d'où, compte tenu de (2),

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx - \int_b^c f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx,$$

c'est-à-dire qu'on retrouve l'égalité (3). ■

3° *Tout facteur constant peut être sorti du signe d'intégration, c'est-à-dire que*

$$\int_a^b kf(x)\,dx = k\int_a^b f(x)\,dx. \tag{4}$$

Démonstration. En effet, pour toute subdivision de $[a, b]$ et tout choix des points ξ_i, on a

$$\sum_{i=1}^n kf(\xi_i)\,\Delta x_i = k\sum_{i=1}^n f(\xi_i)\,\Delta x_i.$$

En faisant tendre λ vers 0, on trouve

$$\int_a^b kf(x)\,dx = \lim_{\lambda\to 0}\sum_{i=1}^n kf(\xi_i)\,\Delta x_i = \lim_{\lambda\to 0} k\sum_{i=1}^n f(\xi_i)\,\Delta x_i =$$

$$= k\lim_{\lambda\to 0}\sum_{i=1}^n f(\xi_i)\,\Delta x_i = k\int_a^b f(x)\,dx,$$

c'est-à-dire l'égalité (4). ■

4° *L'intégrale définie de la somme algébrique de fonctions est égale à la somme algébrique des intégrales de ces fonctions, soit*

$$\int_a^b [f(x) \pm g(x)]\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx.$$

Démonstration. En effet, pour toute subdivision de $[a, b]$ et tout choix des points ξ_i, on a

$$\sum_{i=1}^n [f(\xi_i) \pm g(\xi_i)]\,\Delta x_i = \sum_{i=1}^n f(\xi_i)\,\Delta x_i \pm \sum_{i=1}^n g(\xi_i)\,\Delta x_i.$$

Puisque

$$\lim_{\lambda\to 0}\sum_{i=1}^n f(\xi_i)\,\Delta x_i = \int_a^b f(x)\,dx \quad \text{et} \quad \lim_{\lambda\to 0}\sum_{i=1}^n g(\xi_i)\,\Delta x_i = \int_a^b g(x)\,dx,$$

il vient

$$\int_a^b [f(x) \pm g(x)]\,dx = \lim_{\lambda \to 0} \sum_{i=1}^{n} [f(\xi_i) \pm g(\xi_i)]\,\Delta x_i =$$

$$= \lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i \pm \lim_{\lambda \to 0} \sum_{i=1}^{n} g(\xi_i)\,\Delta x_i = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx. \blacksquare$$

Remarque. La propriété 4° est valable pour tout nombre fini de termes.

§ 5. Estimations des intégrales. Formule de la moyenne

1. Estimations des intégrales. 1° *Si* $f(x) \geqslant 0$ *sur un intervalle* $[a, b]$, *alors*

$$\int_a^b f(x)\,dx \geqslant 0.$$

Démonstration. En effet, toute somme intégrale $\sigma = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i$ de la fonction $f(x)$ sur $[a, b]$ est positive, puisque

$$f(\xi_i) \geqslant 0,\ \Delta x_i = x_i - x_{i-1} > 0,\ i = 1, 2, \ldots, n.$$

En faisant tendre λ vers 0 dans l'inégalité $\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i \geqslant 0$, on obtient

$$\int_a^b f(x)\,dx \geqslant 0. \blacksquare$$

2° *Si* $f(x) \leqslant g(x)$ *sur un intervalle* $[a, b]$, *alors*

$$\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx. \qquad (1)$$

Démonstration. En appliquant l'estimation 1° à la fonction $g(x) - f(x) \geqslant 0$, on obtient

$$\int_a^b [g(x) - f(x)]\,dx \geqslant 0.$$

Or, en vertu de la propriété 4°,

$$\int_a^b [g(x) - f(x)]\,dx = \int_a^b g(x)\,dx - \int_a^b f(x)\,dx \geqslant 0,$$

d'où l'on déduit l'inégalité (1). ■

3° *Pour toute fonction $f(x)$ définie sur un intervalle $[a, b]$ on a*

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx. \tag{2}$$

Démonstration. En appliquant l'estimation 2° à la double inégalité évidente $-|f(x)| \leqslant f(x) \leqslant |f(x)|$, on obtient

$$-\int_a^b |f(x)|\,dx \leqslant \int_a^b f(x)\,dx \leqslant \int_a^b |f(x)|\,dx,$$

ce qui équivaut à (2). ■

Corollaire. *Si $|f(x)| \leqslant k$ sur un intervalle $[a, b]$, alors*

$$\left|\int_a^b f(x)\,dx\right| \leqslant k(b-a). \tag{3}$$

En effet, l'inégalité $|f(x)| \leqslant k$ et les estimations 2° et 3° entraînent

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx \leqslant \int_a^b k\,dx = k\int_a^b dx.$$

En remarquant que

$$\int_a^b dx = \lim_{\lambda\to 0}\sum_{i=1}^{n} 1\cdot\Delta x_i = b-a, \tag{4}$$

on déduit la relation (3).

4° *Si m et M sont respectivement le minimum et le maximum d'une fonction $f(x)$ sur un intervalle $[a, b]$, alors*

$$m(b-a) \leqslant \int_a^b f(x)\,dx \leqslant M(b-a). \tag{5}$$

Démonstration. Par hypothèse, pour tout $x \in [a, b]$, on a

$$m \leqslant f(x) \leqslant M.$$

En appliquant l'estimation 2° à cette double inégalité, on trouve

$$m\int_a^b dx \leqslant \int_a^b f(x)\,dx \leqslant M\int_a^b dx,$$

d'où l'on déduit (5) en tenant compte de (4). ■

2. Formule de la moyenne.

Théorème 8.5 (théorème de la moyenne). *Si une fonction $f(x)$ est continue sur un intervalle $[a, b]$, il existe un point c de $[a, b]$ tel que*

$$\int_a^b f(x)\,dx = f(c)\,(b-a). \tag{6}$$

Démonstration. La fonction $f(x)$ étant continue sue $[a, b]$, le deuxième théorème de Weierstrass affirme l'existence de deux nombres m et M tels que

$$\min_{[a,\,b]} f(x) = m \leqslant f(x) \leqslant M = \max_{[a,\,b]} f(x).$$

Grâce à l'estimation 4° on en déduit que

$$m\,(b-a) \leqslant \int_a^b f(x)\,dx \leqslant M\,(b-a),$$

donc

$$m \leqslant \frac{\int_a^b f(x)\,dx}{b-a} \leqslant M.$$

Posons

$$\frac{\int_a^b f(x)\,dx}{b-a} = \mu \qquad (m \leqslant \mu \leqslant M).$$

Le nombre μ étant compris entre le minimum et le maximum de la fonction $f(x)$ continue sur $[a,b]$ (fig. 98), le théorème 4.10 de passage d'une fonction continue par toute valeur intermédiaire nous dit qu'il existe un point $c \in [a, b]$ tel que $f(c) = \mu$. Donc

$$\frac{\int_a^b f(x)\,dx}{b-a} = f(c),$$

ce qui équivaut à l'égalité (6). ■

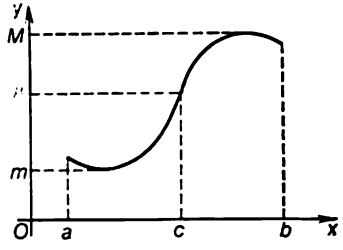


Fig. 98

La formule (6) s'appelle *formule de la moyenne* et la quantité $f(c)$, valeur moyenne de la fonction $f(x)$ sur l'intervalle $[a, b]$.

Remarque. Le théorème de la moyenne admet l'interprétation géométrique suivante : l'intégrale définie d'une fonction $f(x) \geqslant 0$ est égale à l'aire du rectangle de hauteur $f(c)$ et de base $b - a$.

§ 6. Intégrale à borne supérieure variable

Jusqu'alors nous avons étudié l'intégrale définie avec des bornes d'intégration a et b fixes. Si par exemple la borne supérieure d'intégration varie dans les limites de $[a, b]$, l'intégrale prend des valeurs différentes. Autrement dit, une intégrale dont la borne supérieure est variable est une fonction de cette borne.

Considérons l'intégrale $\int_a^x f(t)\,dt$ *) $(a \leqslant x \leqslant b)$ où la borne inférieure a est fixe et la borne supérieure x, variable. Cette intégrale est fonction de sa borne supérieure x. Désignons cette fonction par $\Phi(x)$, c'est-à-dire posons

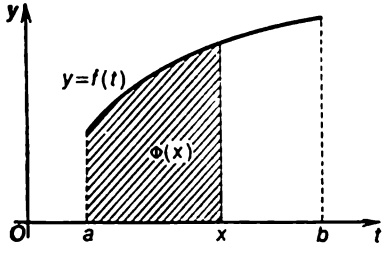


Fig. 99

$$\Phi(x) = \int_a^x f(t)\,dt \qquad (1)$$

et appelons-la *intégrale à borne supérieure variable*. Géométriquement, la fonction $\Phi(x)$ représente l'aire du trapèze curviligne hachuré sur la figure 99.

Le théorème suivant nous donne la valeur d'une intégrale à borne supérieure variable.

Théorème 8.6. *La dérivée d'une intégrale d'une fonction continue par rapport à la borne supérieure d'intégration existe et est égale à la valeur prise par l'intégrant au point égal à la borne supérieure, autrement dit*

$$\Phi'(x) = \left(\int_a^x f(t)\,dt\right)' = f(x). \qquad (2)$$

Démonstration. Considérons une valeur arbitraire $x \in [a, b]$ et donnons-lui un accroissement $\Delta x \neq 0$ tel que $x + \Delta x \in [a, b]$. La fonction $\Phi(x)$ définie par (1) prend alors la valeur

$$\Phi(x+\Delta x) = \int_a^{x+\Delta x} f(t)\,dt.$$

D'après la propriété 2° de l'intégrale définie (cf. § 4), on a

$$\Phi(x+\Delta x) = \int_a^x f(t)\,dt + \int_x^{x+\Delta x} f(t)\,dt = \Phi(x) + \int_x^{x+\Delta x} f(t)\,dt.$$

*) x étant la borne supérieure d'intégration, on a désigné la variable d'intégration par t.

D'où l'on déduit l'accroissement de $\Phi(x)$:

$$\Phi(x+\Delta x)-\Phi(x)=\int_{x}^{x+\Delta x} f(t)\,dt.$$

Le théorème 8.5 nous donne

$$\Phi(x+\Delta x)-\Phi(x)=f(c)\,\Delta x,$$

où c est compris entre x et $x+\Delta x$. Divisons les deux membres de cette égalité par Δx :

$$\frac{\Phi(x+\Delta x)-\Phi(x)}{\Delta x}=f(c).$$

Si maintenant $\Delta x \to 0$, alors $c \to x$ et $f(c) \to f(x)$ en vertu de la continuité de $f(x)$ sur $[a, b]$. Si donc on fait tendre Δx vers 0 dans la dernière égalité, on obtient

$$\lim_{\Delta x\to 0}\frac{\Phi(x+\Delta x)-\Phi(x)}{\Delta x}=\lim_{\Delta x\to 0} f(c)=\lim_{c\to x} f(c)=f(x)$$

où $\Phi'(x)=f(x)$. ■

On a ainsi démontré que toute fonction $f(x)$ continue sur $[a, b]$ possède sur cet intervalle une primitive et de plus la fonction $\Phi(x)$ est une primitive de $f(x)$. Comme toute autre primitive de $f(x)$ diffère de $\Phi(x)$ d'une constante additive (cf. théorème 7.1), on a la relation suivante entre intégrale définie et intégrale indéfinie :

$$\int f(x)\,dx=\int_{a}^{x} f(t)\,dt+C,$$

où C est une constante arbitraire.

§ 7. Formule de Newton-Leibniz

Le calcul des intégrales définies par une méthode basée sur la définition de l'intégrale par une somme intégrale soulève en principe de grosses difficultés. Il existe une méthode de calcul des intégrales définies plus commode qui, on le prouvera plus bas, utilise la relation établie au § 6 entre l'intégrale définie et l'intégrale indéfinie.

On a montré plus haut qu'une fonction $f(x)$ continue sur un intervalle $[a, b]$ possède des primitives sur cet intervalle, l'une d'elles est la fonction

$$\Phi(x)=\int_{a}^{x} f(t)\,dt.$$

Soit $F(x)$ une autre primitive de $f(x)$ sur $[a, b]$. Les primitives $\Phi(x)$ et $F(x)$ différant d'une constante, on a

$$\int_a^x f(t)\,dt = F(x) + C, \quad a \leqslant x \leqslant b,$$

où C est une constante. En faisant $x = a$ dans cette égalité et en se servant de la formule (1) du § 4, on obtient

$$\int_a^a f(t)\,dt = F(a) + C. \quad 0 = F(a) + C, \quad C = -F(a),$$

autrement dit pour tout $x \in [a, b]$ on a

$$\int_a^x f(t)\,dt = F(x) - F(a).$$

En faisant $x = b$, on obtient la formule fondamentale du calcul intégral

$$\int_a^b f(x)\,dx = F(b) - F(a), \tag{1}$$

appelée *formule de Newton-Leibniz.*

On convient de désigner la différence $F(b) - F(a)$ comme suit :

$$F(x)\Big|_a^b \quad \text{ou} \quad [F(x)]_a^b,$$

donc la formule (1) devient

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b.$$

Signalons que dans la formule (1), pour $F(x)$ on peut prendre toute primitive de $f(x)$ sur $[a, b]$.

La formule (1) fournit une méthode simple de calcul de l'intégrale définie : l'intégrale définie d'une fonction continue est égale à la différence des valeurs prises par toute primitive de cette fonction aux bornes supérieure et inférieure. Cette formule offre de riches possibilités pour le calcul des intégrales définies en ramenant ce dernier au calcul de l'intégrale indéfinie qui a été suffisamment bien étudiée.

Considérons des exemples.

1. $\displaystyle\int_a^b \sin x\,dx = -\cos x\Big|_a^b = \cos a - \cos b.$

2. $\int_0^2 (3x^2-1)\,dx = [x^3-x]_0^2 = (2^3-2)-(0^3-0) = 6.$

3. $\int_1^2 \frac{dx}{x} = \ln x\Big|_1^2 = \ln 2 - \ln 1 = \ln 2.$

4. $\int_{-1}^{1} \frac{dx}{1+x^2} = \operatorname{Arc\,tg} x\Big|_{-1}^{1} = \operatorname{Arc\,tg} 1 - \operatorname{Arc\,tg}(-1) =$

$$= \frac{\pi}{4} - \left(-\frac{\pi}{4}\right) = \frac{\pi}{4} + \frac{\pi}{4} = \frac{\pi}{2}.$$

5. $\int_0^3 \frac{dx}{\sqrt{1+x^2}} = \ln\left(x+\sqrt{1+x^2}\right)\Big|_0^3 = \ln(3+\sqrt{10}).$

Remarque. La formule de Newton-Leibniz a été établie sous l'hypothèse que l'intégrant $f(x)$ soit continu. Cette formule a lieu pour les fonctions discontinues moyennant certaines conditions.

§ 8. Changement de variable dans l'intégrale définie

Théorème 8.7. *Soit $f(x)$ une fonction continue sur un intervalle $[a, b]$. Si une fonction $x = \varphi(t)$*

1) *admet une dérivée $\varphi'(t)$ continue sur un intervalle $[\alpha, \beta]$;*
2) *prend ses valeurs sur l'intervalle $[a, b]$;*
3) $\varphi(\alpha) = a$, $\varphi(\beta) = b$ (fig. 100), *alors*

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f[\varphi(t)]\,\varphi'(t)\,dt. \qquad (1)$$

Démonstration. La formule de Newton-Leibniz nous donne

$$\int_a^b f(x)\,dx = F(b) - F(a),$$

où $F(x)$ est une primitive de $f(x)$ sur $[a, b]$. Considérons la fonction composée $\Phi(t) = F[\varphi(t)]$ sur l'intervalle $[\alpha, \beta]$. La règle de dérivation d'une fonction composée nous donne:

$$\Phi'(t) = F'[\varphi(t)]\,\varphi'(t) = f[\varphi(t)]\,\varphi'(t).$$

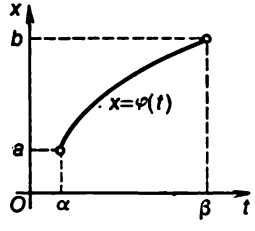


Fig. 100

Il s'ensuit que la fonction $\Phi(t)$ est une primitive de la fonction $f[\varphi(t)]\,\varphi'(t)$ continue sur $[\alpha, \beta]$, donc en vertu de la formule de

Newton-Leibniz,

$$\int_{\alpha}^{\beta} f[\varphi(t)]\,\varphi'(t)\,dt = \Phi(\beta) - \Phi(\alpha) = F[\varphi(\beta)] - F[\varphi(\alpha)] =$$

$$= F(b) - F(a) = \int_{a}^{b} f(x)\,dx.$$

Ce qui prouve la formule (1). ■

La formule (1) s'appelle *formule de changement de la variable* (sous-entendu d'intégration) ou *formule de substitution* dans l'intégrale définie.

Remarque 1. Si dans le calcul d'une intégrale indéfinie par un changement de la variable x en t il faut ensuite revenir à l'ancienne variable x, ceci est inutile dans le calcul d'une intégrale définie, car il faut trouver maintenant un nombre, qui en vertu de la formule prouvée, est égal à la valeur de chaque intégrale envisagée.

Exemple 1. Calculer $\int_{0}^{1} \sqrt{1-x^2}\,dx$.

Solution. Considérons la substitution $x = \sin t$, $0 \leqslant t \leqslant \frac{\pi}{2}$. Vérifions la légitimité de cette substitution.

Primo, la fonction $f(x) = \sqrt{1-x^2}$ est continue sur $[0, 1]$; *secundo*, la fonction $x = \sin t$ est dérivable sur $[0, \pi/2]$ et $x'_t = \cos t$ est continue sur $[0, \pi/2]$; *tertio*, la fonction $x = \sin t$ prend ses valeurs sur l'intervalle $[0, 1]$ et de plus $x(0) = 0$ et $x(\pi/2) = 1$. Cette substitution satisfait donc à toutes les conditions du théorème 8.7. En appliquant la formule (1) on obtient

$$\int_{0}^{1} \sqrt{1-x^2}\,dx = \int_{0}^{\pi/2} \sqrt{1-\sin^2 t}\cos t\,dt = \int_{0}^{\pi/2} \cos^2 t\,dt =$$

$$= \frac{1}{2}\int_{0}^{\pi/2} (1+\cos 2t)\,dt = \frac{\pi}{4}.$$

Remarque. Avant de se servir de la formule (1) il faut bien s'assurer que toutes les conditions du théorème sont remplies. Si ces conditions sont violées, on risque d'obtenir un faux résultat.

Exemple 2. Calculer $\int_{0}^{\pi} dx$.

Solution. On a

$$\int_0^{\pi} dx = x\Big|_0^{\pi} = \pi.$$

Par ailleurs,

$$\int_0^{\pi} dx = \int_0^{\pi} \frac{dx}{\sin^2 x + \cos^2 x} = \int_0^{\pi} \frac{dx}{\cos^2 x\,(1+\operatorname{tg}^2 x)}.$$

La substitution $\operatorname{tg} x = t$ nous conduit formellement au résultat suivant :

$$\int_0^{\pi} dx = \int_0^{\pi} \frac{d(\operatorname{tg} x)}{1+\operatorname{tg}^2 x} = \int_0^{0} \frac{dt}{1+t^2} = 0.$$

Ce résultat est faux, puisque $\pi \neq 0$. La cause en est que la fonction $t = \operatorname{tg} x$ est discontinue pour $x = \pi/2$ et ne satisfait pas aux conditions du théorème 8.7.

§ 9. Formule d'intégration par parties dans l'intégrale définie

Théorème 8.8. *Si $u(x)$ et $v(x)$ sont des fonctions continues sur un intervalle $[a, b]$, alors*

$$\int_a^b u\,dv = uv\Big|_a^b - \int_a^b v\,du. \tag{1}$$

Démonstration. La fonction $u(x)\,v(x)$ étant une primitive de $[u(x)\,v(x)]' = u(x)\,v'(x) + v(x)\,u'(x)$, la formule de Newton-Leibniz nous donne

$$\int_a^b [u(x)\,v'(x) + v(x)\,u'(x)]\,dx = [u(x)\,v(x)]_a^b.$$

En appliquant la propriété 4° des intégrales définies (cf. § 4) on trouve

$$\int_a^b u(x)\,v'(x)\,dx + \int_a^b v(x)\,u'(x)\,dx = [u(x)\,v(x)]_a^b,$$

ou, ce qui est équivalent,

$$\int_a^b u\,dv + \int_a^b v\,du = uv\Big|_a^b,$$

d'où l'on déduit la formule (1). ■

La formule (1) s'appelle *formule d'intégration par parties* dans l'intégrale définie.

Exemple 1. Calculer $\int_1^e \ln x\,dx$.

Solution. Posons $u=\ln x$, $dv=dx$; d'où $du=\frac{dx}{x}$, $v=x$ et la formule (1) nous donne

$$\int_1^e \ln x\,dx = x\ln x\Big|_1^e - \int_1^e x\,\frac{dx}{x} = [x\ln x - x]_1^e = 1.$$

Exemple 2. Calculer $\int_1^2 xe^x\,dx$.

Solution. Posons $u=x$, $dv=e^x\,dx$; d'où $du=dx$, $v=e^x$ et d'après la formule (1)

$$\int_1^2 xe^x\,dx = xe^x|_1^2 - \int_1^2 e^x\,dx = [e^x(x-1)]_1^2 = e^2.$$

Exemple 3. Calculer $\int_0^1 \operatorname{Arc\,tg} x\,dx$.

Solution. Posons $u=\operatorname{Arc\,tg} x$, $dv=dx$; d'où $du=\frac{dx}{1+x^2}$, $v=x$ et la formule (1) nous donne

$$\int_0^1 \operatorname{Arc\,tg} x\,dx = x\operatorname{Arc\,tg} x\Big|_0^1 - \int_0^1 \frac{x\,dx}{1+x^2} =$$

$$= \left[x\operatorname{Arc\,tg} x - \frac{1}{2}\ln(1+x^2)\right]_0^1 = \frac{\pi}{4} - \ln\sqrt{2}.$$

§ 10. Quelques applications physiques et géométriques de l'intégrale définie

1. Aire d'un trapèze curviligne. Soit donnée dans le plan Oxy une figure limitée par un intervalle $[a, b]$ de l'axe Ox, les droites $x=a$, $x=b$ et la courbe représentative d'une fonction $y=f(x)$ continue et positive sur $[a, b]$. Cette figure est un trapèze curviligne dont l'aire s peut être calculée à l'aide de la formule

$$s = \int_a^b f(x)\,dx. \qquad (1)$$

Démonstration. Divisons arbitrairement l'intervalle $[a, b]$ en n parties par les points $a=x_0<x_1<x_2<\ldots<x_{i-1}<x_i<\ldots<x_n=b$, prenons un point ξ_i dans chaque

intervalle partiel $[x_{i-1}, x_i]$ et considérons la figure étagée (fig. 101). L'aire s du trapèze curviligne est approximativement égale à celle de la figure étagée :

$$s \approx \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i,$$

où $\Delta x_i = x_i - x_{i-1}$. Il est naturel d'admettre que l'aire de la figure étagée tend vers celle du trapèze curviligne lorsque $\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta x_i\}$ tend vers 0. Par ailleurs, l'aire de la figure étagée est la somme intégrale de l'intégrale (1). La fonction $f(x)$ étant continue sur $[a, b]$, cette somme admet lorsque $\lambda \to 0$ une limite qui est égale à l'intégrale définie de la fonction $f(x)$ sur $[a, b]$. Donc l'aire s du trapèze curviligne est numériquement égale à l'intégrale définie de la fonction $f(x)$ sur $[a, b]$:

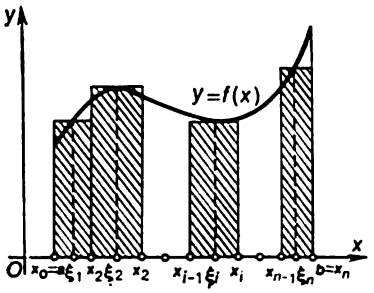


Fig. 101

$$s = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = \int_a^b f(x)\,dx. \quad \blacksquare$$

Ainsi *l'intégrale définie d'une fonction $f(x)$ continue et positive sur $[a, b]$ est numériquement égale à l'aire du trapèze curviligne de base $[a, b]$ limité par la courbe représentative de la fonction $y = f(x)$. Ceci est la signification géométrique de l'intégrale définie.*

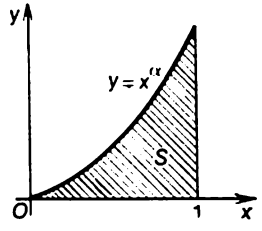


Fig. 102

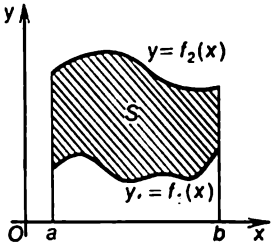


Fig. 103

Exemple 1. Trouver l'aire de la figure limitée par la courbe $y = x^\alpha$, $\alpha > 0$, la droite $x = 1$ et l'axe Ox (fig. 102).

Solution. La formule (1) nous donne

$$s = \int_0^1 x^\alpha\,dx = \frac{x^{\alpha+1}}{\alpha+1}\Big|_0^1 = \frac{1}{\alpha+1}.$$

Si $\alpha = 1$, alors $s = 1/2$; si $\alpha = 2$, $s = 1/3$, etc.

Supposons que la figure est limitée en bas et en haut par des courbes $y = f_1(x)$ et $y = f_2(x)$, $f_1(x) \leqslant f_2(x)$, $a \leqslant x \leqslant b$ (fig. 103),

où $f_1(x)$ et $f_2(x)$ sont des fonctions continues. Si les deux fonctions sont positives, l'aire s de cette figure est égale à la différence des aires des trapèzes curvilignes limités par les courbes $y = f_2(x)$ et $y = f_1(x)$. Donc

$$s = \int_a^b f_2(x)\,dx - \int_a^b f_1(x)\,dx = \int_a^b [f_2(x) - f_1(x)]\,dx. \qquad (2)$$

Remarquons que la formule (2) est valable aussi lorsque $f_1(x)$ et $f_2(x)$ ne sont pas positives. En effet, ces fonctions étant bornées, il existe un nombre $h > 0$ tel que les fonctions $f_1^*(x) = f_1(x) + h$, $f_2^*(x) = f_2(x) + h$ sont positives et l'on a l'égalité évidente

$$\int_a^b [f_2^*(x) - f_1^*(x)]\,dx = \int_a^b [f_2(x) - f_1(x)]\,dx.$$

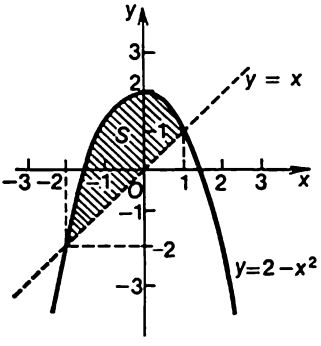


Fig. 104

Exemple 2. Calculer l'aire de la figure limitée par les courbes $y = f_1(x) = x$ et $y = f_2(x) = 2 - x^2$ (fig. 104).

Solution. Trouvons les abscisses des points d'intersection de la droite $y = x$ et de la parabole $y = 2 - x^2$. La résolution du système d'équations

$$\begin{cases} y = x, \\ y = 2 - x^2 \end{cases}$$

nous donne $x_1 = -2$, $x_2 = 1$. Ces valeurs sont les bornes d'intégration. D'après la formule (2) l'aire cherchée est :

$$s = \int_{-2}^{1} [f_2(x) - f_1(x)]\,dx = \int_{-2}^{1} [(2 - x^2) - x]\,dx = \left[2x - \frac{x^3}{3} - \frac{x^2}{2}\right]_{-2}^{1} = \frac{9}{2}.$$

Remarque. Pour calculer l'aire d'un trapèze curviligne dans le cas où la courbe est définie paramétriquement par les équations $x = \varphi(t)$, $y = \psi(t)$, $t \in [\alpha, \beta]$, $\varphi(\alpha) = a$, $\varphi(\beta) = b$, il faut faire le changement de variable $x = \varphi(t)$, $dx = \varphi'(t)\,dt$ dans la formule (1). On obtient alors

$$s = \int_\alpha^\beta \psi(t)\,\varphi'(t)\,dt.$$

Exemple 3. Calculer l'aire de l'ellipse

$$\begin{cases} x = a\cos t, \\ y = b\sin t, \end{cases} \quad 0 \leqslant t \leqslant 2\pi.$$

Solution. L'ellipse étant symétrique par rapport aux axes de coordonnées, il suffit de calculer l'aire de la figure située

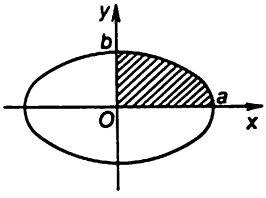


Fig. 105

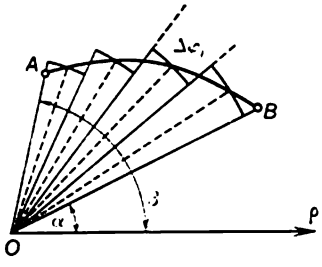


Fig. 106

dans le premier quadrant (fig. 105). L'aire cherchée est donc égale à

$$s = 4\int_{\pi/2}^{0} b\sin t\,(a\cos t)'\,dt = 4ab\int_{0}^{\pi/2}\sin^2 t\,dt =$$

$$= 2ab\int_{0}^{\pi/2}(1-\cos 2t)\,dt = 2ab\left[t-\frac{1}{2}\sin 2t\right]_0^{\pi/2} = \pi ab.$$

Si, en particulier, $a = b = R$, on obtient la formule classique de l'aire du cercle: πR^2.

2. Aire d'un secteur curviligne. Soit AB une courbe d'équation polaire $\rho = \rho(\varphi)$, $\alpha \leqslant \varphi \leqslant \beta$, la fonction $\rho(\varphi)$ étant continue et positive sur $[\alpha, \beta]$. On appellera *secteur curviligne* la figure plane délimitée par la courbe AB et les deux rayons faisant les angles α et β avec l'axe polaire (fig. 106). L'aire s du secteur curviligne est

$$s = \frac{1}{2}\int_{\alpha}^{\beta}\rho^2(\varphi)\,d\varphi. \tag{3}$$

Démonstration. Divisons arbitrairement l'intervalle $[\alpha, \beta]$ en n parties par les points $\alpha = \varphi_0 < \varphi_1 < \varphi_2 < \ldots$ $\ldots < \varphi_{i-1} < \varphi_i < \ldots < \varphi_n = \beta$, choisissons un point arbitraire ξ_i dans chaque intervalle partiel $[\varphi_{i-1}, \varphi_i]$ et construisons des secteurs circulaires de rayons $\rho(\xi_i)$. On obtient une figure en forme d'éventail dont l'aire est approximativement égale à l'aire s

du secteur curviligne:

$$s \approx \frac{1}{2} \sum_{i=1}^{n} \rho^2(\xi_i)\,\Delta\varphi_i,$$

où $\Delta\varphi_i = \varphi_i - \varphi_{i-1}$. L'aire de la figure en éventail est égale à la somme intégrale de l'intégrale (3). La fonction $\rho^2(\varphi)$ étant continue sur l'intervalle $[\alpha, \beta]$, cette somme admet lorsque $\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta\varphi_i\} \to 0$ une limite qui est égale à l'intégrale (3).

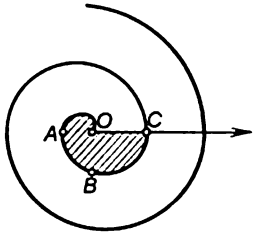


Fig. 107

Donc l'aire du secteur curviligne est numériquement égale à cette intégrale définie:

$$s = \frac{1}{2} \lim_{\lambda \to 0} \sum_{i=1}^{n} \rho^2(\xi_i)\,\Delta\varphi_i = \frac{1}{2} \int_{\alpha}^{\beta} \rho^2(\varphi)\,d\varphi. \blacksquare$$

Exemple 4. Calculer l'aire de la figure délimitée par l'axe polaire et la première spire de la spirale d'Archimède: $\rho = a\varphi$, où $a > 0$ est un nombre (fig. 107).

Solution. Lorsque φ varie entre 0 et 2π, le rayon polaire parcourt la courbe délimitant le secteur $OABC$. La formule (3) nous donne

$$s_{OABC} = \frac{a^2}{2} \int_0^{2\pi} \varphi^2\,d\varphi = \frac{a^2}{2} \frac{\varphi^3}{3}\Big|_0^{2\pi} = \frac{a^2}{2} \frac{8\pi^3}{3} = \frac{4}{3}\pi^3 a^2.$$

La distance du point C au pôle est égale à $\rho = 2\pi a$. Donc l'aire du cercle de rayon OC est $\pi \cdot OC^2 = 4\pi^3 a^2 = 3 \cdot \frac{4}{3}\pi^3 a^2 = 3 s_{OABC}$, c'est-à-dire que l'aire de la figure délimitée par l'axe polaire et la première spire de la spirale d'Archimède est égale à 1/3 de l'aire du cercle de rayon égal au plus grand rayon polaire de la spire. Ce fait a été établi par Archimède.

3. Longueur d'un arc de courbe. Soit AB une courbe plane d'équation $y = f(x)$, $x \in [a, b]$, où $f(x)$ est une fonction continue sur $[a, b]$. Divisons la courbe AB en n parties arbitraires par les points $A = M_0, M_1, \ldots, M_{i-1}, M_i, \ldots, M_n = B$ dans le sens de A vers B. En joignant les points voisins par des cordes, on obtient une ligne polygonale dont on désignera la longueur par P (fig. 108). Désignons par l_i la longueur du segment $M_{i-1}M_i$ de ligne polygonale et par $\mu = \max\limits_{1 \leqslant i \leqslant n} \{l_i\}$.

Définition. On dit qu'un nombre L est la *limite des longueurs P de lignes polygonales lorsque* $\mu \to 0$ ($L = \lim\limits_{\mu \to \infty} P$) si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que

$$| L - P | < \varepsilon$$

pour toute ligne polygonale telle que $\mu < \delta$.

Si cette limite existe, on l'appelle *longueur de l'arc AB.*

Si une fonction $f(x)$ est continue avec sa dérivée $f'(x)$ sur un intervalle $[a, b]$, la longueur L de l'arc AB est donnée par la formule

$$L = \int_a^b \sqrt{1 + f'^2(x)}\, dx. \quad (4)$$

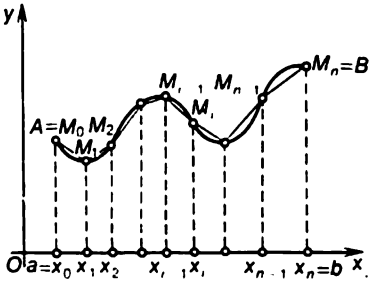


Fig. 108

Démonstration. Désignons par $(x_i, f(x_i))$ les coordonnées du point M_i. Pour les abscisses de ces points on a donc:

$$a = x_0 < x_1 < x_2 < \ldots$$
$$\ldots < x_{i-1} < x_i < \ldots < x_n = b.$$

La longueur l_i d'un segment de ligne polygonale est alors égale à

$$l_i = \sqrt{(x_i - x_{i-1})^2 + [f(x_i) - f(x_{i-1})]^2}.$$

D'après la formule de Lagrange

$$f(x_i) - f(x_{i-1}) = f'(\xi_i)(x_i - x_{i-1}), \quad x_{i-1} \leqslant \xi_i \leqslant x_i.$$

Donc

$$l_i = \sqrt{1 + f'^2(\xi_i)}\, \Delta x_i, \quad \Delta x_i = x_i - x_{i-1}.$$

Et par suite la longueur de la ligne polygonale est

$$P = \sum_{i=1}^{n} l_i = \sum_{i=1}^{n} \sqrt{1 + f'^2(\xi_i)}\, \Delta x_i.$$

Le second membre est la somme intégrale de l'intégrale (4). La fonction $\sqrt{1 + f'^2(x)}$ étant continue sur $[a, b]$, cette somme admet lorsque $\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta x_i\} \to 0$ une limite qui est égale à l'intégrale définie (4). Puisque $\lambda \leqslant \mu$*), on a $\lambda \to 0$ lorsque $\mu \to 0$. Donc

$$L = \lim_{\mu \to 0} P = \lim_{\lambda \to 0} \sum_{i=1}^{n} \sqrt{1 + f'^2(\xi_i)}\, \Delta x_i = \int_a^b \sqrt{1 + f'^2(x)}\, dx. \blacksquare$$

*) $l_i = \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}$, d'où $| \Delta x_i | \leqslant l_i$.

Exemple 5. Rectifier l'arc de parabole semi-cubique $y = x^{3/2}$ pour $x \in [0, 5]$ (fig. 109).

Solution. On a $y' = \frac{3}{2} x^{1/2}$. Donc la formule (4) nous donne

$$L = \int_0^5 \sqrt{1 + y'^2}\, dx = \int_0^5 \sqrt{1 + \frac{9x}{4}}\, dx = \frac{8}{27} \left(1 + \frac{9x}{4}\right)^{3/2} \Big|_0^5 = \frac{335}{27}.$$

Remarque 1. Pour rectifier un arc dans le cas où la courbe AB est définie par des équations paramétriques $x = \varphi(t)$, $y = \psi(t)$,

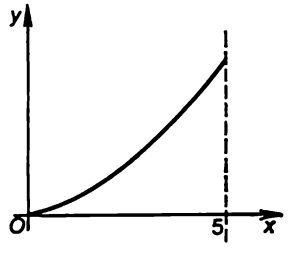


Fig. 109

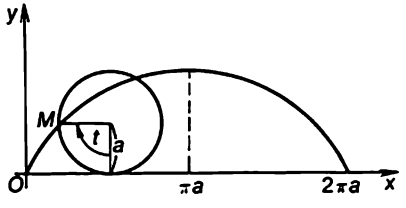


Fig. 110

$\alpha \leqslant t \leqslant \beta$, où α et β sont les valeurs du paramètre t correspondant à $x = a$ et $x = b$, c'est-à-dire $a = \varphi(\alpha)$, $b = \varphi(\beta)$, il faut effectuer le changement $x = \varphi(t)$, $dx = \varphi'(t)\, dt$ dans la formule $L = \int_a^b \sqrt{1 + y'^2(x)}\, dx$. On obtient alors

$$L = \int_a^b \sqrt{1 + y'^2(x)}\, dx = \int_\alpha^\beta \sqrt{1 + \left(\frac{\psi'(t)}{\varphi'(t)}\right)^2} \varphi'(t)\, dt =$$

$$= \int_\alpha^\beta \sqrt{\varphi'^2(t) + \psi'^2(t)}\, dt. \qquad (5)$$

Exemple 6. Rectifier une arche de cycloïde *) $x = a(t - \sin t)$, $y = a(1 - \cos t)$, $t \in [0, 2\pi]$ (fig. 110).

Solution. On a $\varphi'(t) = a(1 - \cos t)$, $\psi'(t) = a \sin t$. Lorsque x parcourt l'intervalle $[0, 2\pi a]$, le paramètre t parcourt

*) La cycloïde est la trajectoire d'un point M d'un cercle de rayon a roulant sans glisser sur une droite.

l'intervalle $[0, 2\pi]$. La longueur cherchée est donc

$$L = \int_0^{2\pi a} \sqrt{1 + y'^2(x)}\, dx = \int_0^{2\pi} \sqrt{\varphi'^2(t) + \psi'^2(t)}\, dt =$$

$$= \int_0^{2\pi} a \sqrt{(1 - \cos t)^2 + \sin^2 t}\, dt = 2a \int_0^{2\pi} \sin \frac{t}{2}\, dt =$$

$$= -4a \cos \frac{t}{2} \Big|_0^{2\pi} = 8a.$$

Remarque 2. Pour rectifier un arc dans le cas où la courbe AB est définie par une équation polaire $\rho = \rho(\varphi)$, $\varphi \in [\alpha, \beta]$, où $\rho(\varphi)$ admet une dérivée $\rho'(\varphi)$ continue sur $[\alpha, \beta]$ et les points A et B correspondent aux valeurs α et β de φ, il faut passer des coordonnées polaires [cf. chap. 3, § 3, formule (1)] aux coordonnées cartésiennes. La courbe AB est alors définie par les équations paramétriques $x = \rho(\varphi) \cos \varphi$, $y = \rho(\varphi) \sin \varphi$, $\varphi \in [\alpha, \beta]$ (φ est le paramètre). Comme

$$x'(\varphi) = \rho'(\varphi) \cos \varphi - \rho(\varphi) \sin \varphi,$$

$$y'(\varphi) = \rho'(\varphi) \sin \varphi + \rho(\varphi) \cos \varphi,$$

la formule (5) devient

$$L = \int_\alpha^\beta \sqrt{\rho(\varphi) + \rho'^2(\varphi)}\, d\varphi. \tag{6}$$

Exemple 7. Rectifier la première spire de la spirale d'Archimède: $\rho = a\varphi$ (fig. 107).

Solution. L'angle φ varie entre 0 et 2π. Donc d'après la formule (6)

$$L = \int_0^{2\pi} \sqrt{a^2\varphi^2 + a^2}\, d\varphi = a \int_0^{2\pi} \sqrt{\varphi^2 + 1}\, d\varphi =$$

$$= a\left[\pi \sqrt{4\pi^2 + 1} + \frac{1}{2} \ln (2\pi + \sqrt{4\pi^2 + 1})\right].$$

4. Volume d'un solide de révolution. Soit $f(x)$ une fonction continue et positive sur un intervalle $[a, b]$. Le volume du solide engendré par la rotation autour de l'axe Ox du trapèze curviligne limité par le graphique de la fonction $y = f(x)$ est:

$$v = \pi \int_a^b f^2(x)\, dx \tag{7}$$

Démonstration. Divisons arbitrairement l'intervalle $[a, b]$ en n parties par les points $a = x_0 < x_1 < x_2 < \ldots$

$\ldots < x_{i-1} < x_i < \ldots < x_n = b$. Construisons un rectangle sur chaque intervalle partiel $[x_{i-1}, x_i]$ (fig. 111). Chaque rectangle engendre un cylindre en tournant autour de l'axe Ox. Trouvons le volume du cylindre i engendré par la rotation du rectangle $PMNQ$:

$$v_i = \pi f^2(x_{i-1})\,\Delta x_i,$$

où $\Delta x_i = x_i - x_{i-1}$.

La somme des volumes des n cylindres est approximativement égale au volume du solide de révolution engendré:

$$v \approx \sum_{i=1}^{n} \pi f^2(x_{i-1})\,\Delta x_i.$$

Par ailleurs, cette somme est la somme intégrale de l'intégrale (7). La fonction $f^2(x)$ étant continue sur $[a, b]$, elle admet lorsque

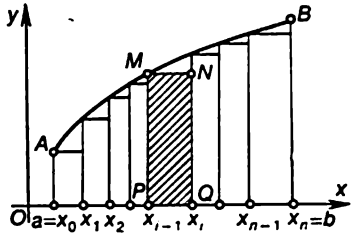


Fig. 111

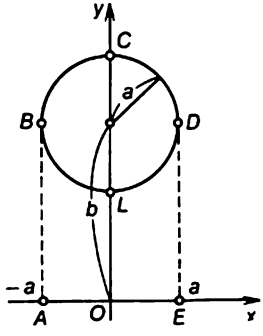


Fig. 112

$\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta x_i\} \to 0$ une limite égale à l'intégrale définie (7). Donc

$$v = \lim_{\lambda \to 0} \sum_{i=1}^{n} \pi f^2(x_{i-1})\,\Delta x_i = \pi \int_a^b f^2(x)\,dx. \quad \blacksquare$$

Exemple 8. Calculer le volume du tore. (Le tore est le solide engendré par un disque de rayon a tournant autour d'un axe contenu dans le même plan et situé à une distance $b > a$ de son centre.)

Solution. Supposons que le disque tourne autour de l'axe Ox (fig. 112). On peut se représenter le volume du tore comme la différence des volumes engendrés par les trapèzes curvilignes $ABCDE$ et $ABLDE$ en tournant autour de l'axe Ox.

Le cercle $LBCD$ a pour équation

$$x^2 + (y - b)^2 = a^2,$$

la courbe BCD,

$$y = y_1(x) = b + \sqrt{a^2 - x^2},$$

la courbe BLD,

$$y = y_2(x) = b - \sqrt{a^2 - x^2}.$$

La formule (7) nous donne le volume v cherché:

$$v = 2\pi \int_0^a y_1^2\,dx - 2\pi \int_0^a y_2^2\,dx = 2\pi \int_0^a (y_1^2 - y_2^2)\,dx =$$

$$= 2\pi \int_0^a [(b + \sqrt{a^2 - x^2})^2 - (b - \sqrt{a^2 - x^2})^2]\,dx =$$

$$= 8\pi b \int_0^a \sqrt{a^2 - x^2}\,dx = 2\pi^2 a^2 b.$$

5. Aire d'une surface de révolution. Soit $f(x)$ une fonction positive et continue avec sa dérivée première sur un intervalle $[a, b]$. L'aire P de la surface engendrée par la courbe représentative de $f(x)$ en tournant autour de l'axe Ox peut être calculée par la formule

$$P = 2\pi \int_a^b f(x) \sqrt{1 + f'^2(x)}\,dx. \quad (8)$$

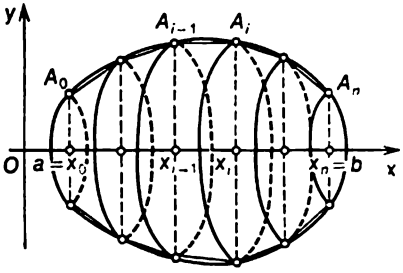


Fig. 113

Démonstration. Divisons arbitrairement l'intervalle $[a, b]$ en n parties par les points $a = x_0 < x_1 < x_2 < \ldots < x_{i-1} < x_i < \ldots < x_n = b$. Soient A_0, A_1, A_2, $\ldots$, A_{i-1}, A_i, $\ldots$, A_n les points correspondants de la courbe. Construisons la ligne polygonale $A_0, A_1, A_2, \ldots, A_n$ (fig. 113). En tournant autour de l'axe Ox, cette ligne engendre une surface composée des surfaces latérales de troncs de cône (cylindre). L'aire de la surface latérale d'un tronc de cône (cylindre) engendré par la révolution du segment i de la ligne polygonale est égale à $2\pi \frac{f(x_{i-1}) + f(x_i)}{2} l_i$ où l_i est la longueur de la corde $A_{i-1}A_i$, c'est-à-dire

$$l_i = \sqrt{(x_i - x_{i-1})^2 + [f(x_i) - f(x_{i-1})]^2}.$$

La formule de Lagrange nous donne

$$f(x_i) - f(x_{i-1}) = f'(\xi_i)(x_i - x_{i-1}), \qquad x_{i-1} \leqslant \xi_i \leqslant x_i.$$

En posant $x_i - x_{i-1} = \Delta x_i$, on trouve

$$l_i = \sqrt{1 + f'^2(\xi_i)}\,\Delta x_i.$$

Ainsi l'aire P de la surface de révolution est approximativement égale à l'aire de la surface engendrée par la révolution de la ligne polygonale :

$$P \approx \sum_{i=1}^{n} 2\pi \frac{f(x_{i-1}) + f(x_i)}{2} \sqrt{1 + f'^2(\xi_i)}\,\Delta x_i.$$

Représentons cette expression par les deux sommes suivantes :

$$P \approx 2\pi \sum_{i=1}^{n} f(\xi_i) \sqrt{1 + f'^2(\xi_i)}\,\Delta x_i +$$

$$+ \pi \left\{ \sum_{i=1}^{n} [(f(x_{i-1}) - f(\xi_i)) + (f(x_i) - f(\xi_i))] \sqrt{1 + f'^2(\xi_i)}\,\Delta x_i \right\}. \quad (9)$$

La première somme du second membre est la somme intégrale de l'intégrale (8), puisque $f(x)\sqrt{1 + f'^2(x)}$ est continue et $\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta x_i\} \to 0$. Montrons que l'expression entre accolades tend vers 0 avec λ. En effet, la fonction $f(x)$ étant uniformément continue sur $[a, b]$, le théorème de Cantor nous dit que pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que $|f(x_{i-1}) - f(\xi_i)| < \varepsilon$ et $|f(x_i) - f(\xi_i)| < \varepsilon$ pour $\lambda < \delta$. Si l'on désigne par M le maximum de la fonction $\sqrt{1 + f'^2(x)}$ sur l'intervalle $[a, b]$, l'expression entre accolades admet la majoration suivante pour $\lambda < \delta$:

$$\left| \left\{ \sum_{i=1}^{n} [(f(x_{i-1}) - f(\xi_i)) + (f(x_i) - f(\xi_i))] \sqrt{1 + f'^2(\xi_i)}\,\Delta x_i \right| <$$

$$< 2M\varepsilon \sum_{i=1}^{n} \Delta x_i = 2M(b - a)\varepsilon.$$

Le nombre ε étant arbitrairement petit, il s'ensuit que cette expression tend vers 0 avec λ.

Donc, en faisant tendre λ vers 0 dans (9), on obtient

$$P = 2\pi \int_a^b f(x) \sqrt{1 + f'^2(x)}\,dx,$$

c'est-à-dire la formule (8). ■

Remarque. Si la surface est engendrée par la révolution autour de l'axe Ox d'une courbe AB d'équations paramétriques $x = \varphi(t)$, $y = \psi(t)$, $t \in [\alpha, \beta]$, où $\psi(t) \geqslant 0$, $\varphi(t)$ varie entre a et b lorsque t varie entre α et β, $\varphi(\alpha) = a$, $\varphi(\beta) = b$, alors la subs-

titution $x = \varphi(t)$ dans l'intégrale (8) nous donne

$$P = 2\pi \int_{\alpha}^{\beta} \psi(t) \sqrt{\varphi_{(t)}'^2 + \psi'^2(t)}\, dt. \tag{10}$$

Si enfin la courbe est définie en coordonnées polaires par l'équation $\rho = \rho(\varphi)$, $\varphi \in [\alpha, \beta]$, où $\rho(\varphi)$ admet une dérivée continue sur $[\alpha, \beta]$, on se ramène, comme indiqué au n° 3, à une courbe définie par les équations paramétriques $x = \rho(\varphi) \cos \varphi$, $y = \rho(\varphi) \sin \varphi$, $\varphi \in [\alpha, \beta]$, et la formule (10) devient

$$P = 2\pi \int_{\alpha}^{\beta} \rho(\varphi) \sin \varphi \sqrt{\rho^2(\varphi) + \rho'^2(\varphi)}\, d\varphi.$$

Exemple 9. Calculer l'aire P de la zone sphérique engendrée par le demi-cercle $f(x) = \sqrt{R^2 - x^2}$, $-R < a \leqslant x \leqslant b < R$ en tournant autour de l'axe Ox.

Solution. La formule (8) nous donne

$$P = 2\pi \int_{a}^{b} \sqrt{R^2 - x^2} \sqrt{1 + \frac{x^2}{R^2 - x^2}}\, dx = 2\pi \int_{a}^{b} R\, dx =$$

$$= 2\pi R (b - a) = 2\pi R h,$$

où h est la hauteur de la zone.

Exemple 10. Calculer l'aire P de la surface engendrée par une arche de cycloïde $x = a(t - \sin t)$, $y = a(1 - \cos t)$, $t \in [0, 2\pi]$, tournant autour de l'axe Ox.

Solution. La formule (10) nous donne

$$P = 2\pi \int_{0}^{2\pi} a(1 - \cos t) \sqrt{(a \sin t)^2 + (a(1 - \cos t))^2}\, dt =$$

$$= 2\sqrt{2}\pi a^2 \int_{0}^{2\pi} (1 - \cos t)^{3/2}\, dt = \frac{64}{3}\pi a^2.$$

6. Travail d'une force variable. Les problèmes liés aux applications géométriques de l'intégrale définie, qui ont été envisagés plus haut, ont été résolus par la même méthode: la quantité cherchée a été représentée approximativement par une somme intégrale, puis sa valeur exacte a été déterminée sous forme d'intégrale par passage à la limite. Cette méthode permet d'aborder de nombreux autres problèmes de mécanique, de physique et de technique. Calculons à titre d'exemple le travail d'une force variable.

Supposons qu'un point matériel se déplace sous l'action d'une force variable dirigée suivant l'axe Ox et dépendant de x. On

demande de déterminer le travail A fourni par la force F lorsque le point matériel se déplace sur l'axe Ox d'un point $x = a$ à un point $x = b$ $(a < b)$. La fonction $F(x)$ est supposée continue sur l'intervalle $[a, b]$ (fig. 114).

Divisons arbitrairement l'intervalle $[a, b]$ en n parties par les points $a = x_0 < x_1 < x_2 < \ldots < x_{i-1} < x_i < \ldots < x_n = b$. Choisissons dans chaque intervalle partiel $[x_{i-1}, x_i]$ un point arbitraire ξ_i. La force qui agit sur le point matériel varie d'un point à l'autre de l'intervalle $[x_{i-1}, x_i]$. Mais si l'intervalle

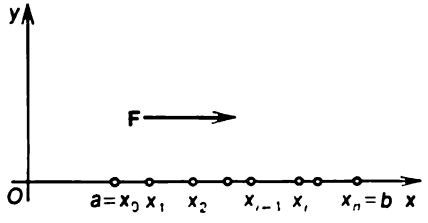


Fig. 114

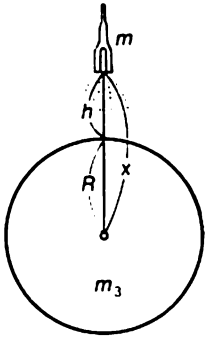


Fig. 115

$[x_{i-1}, x_i]$ est petit, la force prend sur cet intervalle des valeurs qui diffèrent peu de celle prise en ξ_i, puisque F est continue. Donc on peut considérer que le travail A_i de la force F sur $[x_{i-1}, x_i]$ est approximativement égal au travail d'une force constante $F(\xi_i)$ sur le même intervalle, soit

$$A_i \approx F(\xi_i)\,\Delta x_i.$$

En raisonnant de même pour chaque intervalle de la subdivision, on trouve que la valeur approchée du travail A de la force F sur l'intervalle tout entier est

$$A \approx \sum_{i=1}^{n} F(\xi_i)\,\Delta x_i.$$

La somme du second membre est la somme intégrale de la fonction $F(x)$. La fonction $F(x)$ étant continue sur $[a, b]$, cette somme possède lorsque $\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta x_i\} \to 0$ une limite égale à l'intégrale définie de la fonction $F(x)$ sur l'intervalle $[a, b]$. Donc

$$A = \lim_{\lambda \to 0} \sum_{i=1}^{n} F(\xi_i)\,\Delta x_i = \int_a^b F(x)\,dx. \tag{11}$$

Exemple 11. Déterminer le travail A nécessaire pour lancer un solide de masse m à une altitude h (fig. 115).

Solution. Désignons par F la force d'attraction de la Terre. Soit m_t la masse de la Terre. D'après la loi de Newton

$$F = G\frac{mm_t}{x^2},$$

où x est la distance du solide au centre de la Terre. En posant $Gmm_t = K$, on obtient $F(x) = K/x^2$, $R \leqslant x \leqslant h + R$, où R est le rayon de la Terre. Pour $x = R$ la force $F(R)$ est égale au poids $P = mg$ du solide, c'est-à-dire que $K/R^2 = P$, d'où $K = PR^2$ et $F(x) = PR^2/x^2$. La formule (11) nous donne donc

$$A = \int_R^{R+h} F(x)\,dx = PR^2 \int_R^{R+h} \frac{dx}{x^2} = -PR^2 \frac{1}{x}\Big|_R^{R+h} = \frac{PRh}{R+h}.$$

§ 11. Intégrales impropres

En introduisant l'intégrale définie comme la limite de sommes intégrales, nous avons admis que l'intervalle d'intégration était fini et l'intégrant, borné sur cet intervalle. Si l'une au moins de ces conditions n'est pas remplie, la définition de l'intégrale n'a plus de sens. C'est ainsi que si l'intervalle d'intégration est infini, on ne peut le subdiviser en n parties de longueur finie et si l'intégrant est infini, sa somme intégrale ne peut tendre vers une limite finie. Il n'empêche toutefois que la notion d'intégrale définie peut être étendue à ces cas. Cette généralisation nous conduit à la notion d'intégrale impropre.

1. Intégrales impropres à bornes d'intégration infinies.

Définition. Soit $f(x)$ une fonction définie sur un intervalle $[a, +\infty[$ et intégrable sur tout intervalle $[a, R]$, c'est-à-dire qu'existe l'intégrale définie $\int_a^R f(x)\,dx$ pour tout $R > a$. La limite finie

$$\lim_{R\to+\infty} \int_a^R f(x)\,dx, \tag{1}$$

si elle existe, s'appelle *intégrale impropre de première espèce* et se note

$$\int_a^{+\infty} f(x)\,dx. \tag{2}$$

Ainsi par définition

$$\int_a^{+\infty} f(x)\,dx = \lim_{R\to+\infty} \int_a^R f(x)\,dx.$$

Dans ce cas on dit que l'intégrale (2) existe ou converge. Si la limite (1) n'existe pas ou est infinie, on dit que l'intégrale (2) n'existe pas ou diverge.

Par analogie à l'intégrale (2), on envisage l'intégrale impropre sur l'intervalle $]-\infty, b]$:

$$\int_{-\infty}^{b} f(x)\,dx = \lim_{R\to-\infty} \int_{R}^{b} f(x)\,dx. \tag{3}$$

Enfin l'intégrale impropre à bornes infinies peut être définie par la somme d'intégrales (2) et (3), c'est-à-dire que

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{c} f(x)\,dx + \int_{c}^{+\infty} f(x)\,dx, \tag{4}$$

où c est un nombre arbitraire, sous réserve que les deux intégrales du second membre existent.

Etablissons la signification géométrique de l'intégrale impropre de première espèce. Soit $f(x) \geqslant 0$. L'intégrale définie

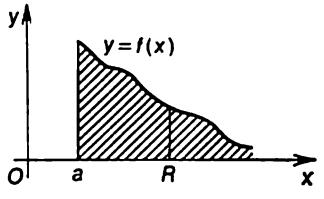


Fig. 116

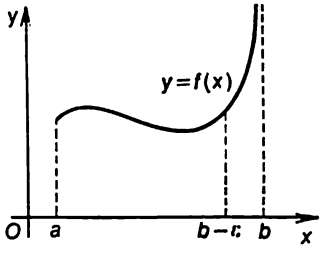


Fig. 117

$\int_{a}^{R} f(x)\,dx$ exprime alors l'aire du domaine délimité par la courbe représentative de la fonction $f(x)$, l'axe Ox, les droites $x = a$ et $x = R$. On admet naturellement que l'intégrale impropre $\int_{a}^{+\infty} f(x)\,dx$ exprime l'aire finie d'un domaine infini délimité par la courbe représentative de $f(x)$, l'axe Ox et la droite $x = a$ (fig. 116). Les intégrales (3) et (4) admettent une interprétation analogue.

Voyons quelques exemples de calcul des intégrales impropres de première espèce.

Exemple 1. $\int_0^{+\infty} \frac{dx}{1+x^2} = \lim_{R\to+\infty} \int_0^R \frac{dx}{1+x^2} = \lim_{R\to+\infty} \operatorname{Arc\,tg} x \,|_0^R =$

$= \lim_{R\to+\infty} \operatorname{Arc\,tg} R = \frac{\pi}{2}$,

c'est-à-dire que cette intégrale converge.

Exemple 2. $\int_0^{+\infty} \cos x\, dx = \lim_{R\to+\infty} \int_0^R \cos x\, dx = \lim_{R\to+\infty} \sin x \,|_0^R =$

$= \lim_{R\to+\infty} \sin R.$

Cette intégrale diverge puisque la fonction sin R ne possède pas de limite lorsque $R \to +\infty$.

Exemple 3. $\int_{-\infty}^{+\infty} e^x\, dx = \int_{-\infty}^{0} e^x\, dx + \int_0^{+\infty} e^x\, dx.$

Cette intégrale diverge, puisque

$$\int_0^{+\infty} e^x\, dx = \lim_{R\to+\infty} \int_0^R e^x\, dx = \lim_{R\to+\infty} e^x \,|_0^R = \lim_{R\to+\infty} (e^R - 1) = \infty.$$

Exemple 4. $\int_1^{+\infty} \frac{dx}{x^\alpha}$, où α est un nombre.

1) Si $\alpha \neq 1$, on a pour tout $R > 0$

$$\lim_{R\to+\infty} \int_1^R \frac{dx}{x^\alpha} = \lim_{R\to+\infty} \frac{x^{1-\alpha}}{1-\alpha}\Big|_1^R = \lim_{R\to+\infty} \frac{R^{1-\alpha}-1}{1-\alpha} = \begin{cases} \frac{-1}{1-\alpha} & \text{pour } \alpha > 1, \\ \infty & \text{pour } \alpha < 1. \end{cases}$$

2) Si $\alpha = 1$, on a pour tout $R > 0$

$$\lim_{R\to+\infty} \int_1^R \frac{dx}{x} = \lim_{R\to+\infty} \ln x \,|_1^R = \lim_{R\to+\infty} \ln R = \infty.$$

Donc cette intégrale converge pour $\alpha > 1$ et diverge pour $\alpha \leqslant 1$.

On remarquera que le calcul de ces intégrales impropres est basé sur leur définition.

2. Intégrales impropres de fonctions illimitées.

Définition. Soit $f(x)$ une fonction définie sur un intervalle $[a, b[$. On dit que le point $x = b$ est *singulier* si la fonction $f(x)$ est infinie dans tout voisinage de ce point et bornée sur tout intervalle $[a, b - \varepsilon]$ contenu dans $[a, b[$ (fig. 117). Supposons que la fonction $f(x)$ est intégrable sur tout intervalle $[a, b - \varepsilon]$, c'est-à-dire qu'existe l'intégrale définie $\int_a^{b-\varepsilon} f(x)\, dx$ pour tout

$\varepsilon > 0$ tel que $b - \varepsilon > a$. Alors la limite finie

$$\lim_{\varepsilon \to 0+} \int_a^{b-\varepsilon} f(x)\,dx, \tag{5}$$

si elle existe, s'appelle *intégrale impropre de seconde espèce* et se note

$$\int_a^b f(x)\,dx. \tag{6}$$

On dit alors que l'intégrale (6) existe ou converge. Si la limite (5) n'existe pas ou est infinie, on dit que l'intégrale n'existe pas ou diverge.

De façon analogue, si $x = a$ est un point singulier, l'intégrale impropre se définit de la manière suivante:

$$\int_a^b f(x)\,dx = \lim_{\varepsilon \to 0+} \int_{a+\varepsilon}^b f(x)\,dx.$$

Si la fonction $f(x)$ est infinie au voisinage d'un point intérieur $c \in]a, b[$, on convient par définition

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx,$$

sous réserve qu'existent les deux intégrales du second membre.

Enfin, si a et b sont des points singuliers et si les deux intégrales du second membre existent, l'intégrale impropre se définit comme la somme

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx,$$

où c est un point quelconque de $]a, b[$.

Exemple 5. $\int_0^1 \frac{dx}{x^\alpha}$, $\alpha > 0$ est un nombre quelconque.

1) Si $\alpha \neq 1$, on a

$$\lim_{\varepsilon \to 0+} \int_\varepsilon^1 \frac{dx}{x^\alpha} = \lim_{\varepsilon \to 0+} \frac{x^{1-\alpha}}{1-\alpha}\Big|_\varepsilon^1 = \lim_{\varepsilon \to 0+} \frac{1-\varepsilon^{1-\alpha}}{1-\alpha} =$$

$$= \begin{cases} \infty & \text{pour } \alpha > 1, \\ \dfrac{1}{1-\alpha} & \text{pour } 0 < \alpha < 1. \end{cases}$$

2) Si $\alpha = 1$, alors

$$\lim_{\varepsilon \to 0+} \int_{\varepsilon}^{1} \frac{dx}{x} = \lim_{\varepsilon \to 0+} \ln x \Big|_{\varepsilon}^{1} = \lim_{\varepsilon \to 0+} (-\ln \varepsilon) = \infty.$$

Donc cette intégrale converge pour $0 < \alpha < 1$ et diverge pour $\alpha \geqslant 1$.

3. Critère de convergence des intégrales impropres. Etudions la convergence des intégrales impropres de la forme $\int_a^{+\infty} f(x)\,dx$.

Théorème 8.9 (critère de comparaison des intégrales impropres). *Si des fonctions $f(x)$ et $g(x)$, $x \in [a, \infty[$, sont continues et telles que $0 \leqslant f(x) \leqslant g(x)$, alors la convergence de l'intégrale*

$$\int_a^{+\infty} g(x)\,dx \tag{7}$$

entraîne celle de l'intégrale

$$\int_a^{+\infty} f(x)\,dx, \tag{8}$$

et la divergence de l'intégrale (8), *celle de l'intégrale* (7).

Démonstration. Posons

$$F(R) = \int_a^R f(x)\,dx, \quad G(R) = \int_a^R g(x)\,dx.$$

Comme $0 \leqslant f(x) \leqslant g(x)$ pour $x \geqslant a$, on a en vertu des estimations 1° et 2° (cf. § 5)

$$0 \leqslant F(R) \leqslant G(R) \quad \text{pour} \quad R \geqslant a, \tag{9}$$

et de plus la fonction $F(R)$ (de même que $G(R)$) est croissante sur l'intervalle $[a, +\infty[$. En effet, si $a \leqslant R_1 \leqslant R_2$, alors $\int_{R_1}^{R_2} f(x)\,dx \geqslant 0$ et par suite

$$F(R_2) = \int_a^{R_2} f(x)\,dx = \int_a^{R_1} f(x)\,dx + \int_{R_1}^{R_2} f(x)\,dx \geqslant \int_a^{R_1} f(x)\,dx = F(R_1).$$

Supposons que l'intégrale (7) converge, c'est-à-dire que la fonction $G(R)$ possède une limite finie lorsque $R \to +\infty$. Il s'ensuit que la fonction $G(R)$ est bornée sur $[a, +\infty[$, puisqu'elle est croissante.

La fonction $F(R)$ est aussi bornée sur $[a, +\infty[$ en vertu de (9) et par suite elle admet un suprémum sur cet intervalle. Soit $\sup_{[a, +\infty[} F(R) = I$. Par définition du suprémum, pour tout $\varepsilon > 0$, il existe un $R_\varepsilon \geqslant a$ tel que $0 \leqslant I - F(R_\varepsilon) < \varepsilon$. La fonction $F(R)$ étant croissante sur $[a, +\infty[$, on a $F(R) \geqslant F(R_\varepsilon)$ pour tout $R \geqslant R_\varepsilon$ et par suite $0 \leqslant I - F(R) < \varepsilon$ pour $R \geqslant R_\varepsilon$.

Donc

$$|F(R) - I| < \varepsilon \quad \text{pour} \quad R \geqslant R_\varepsilon.$$

Ce qui exprime que $\lim_{R \to +\infty} F(R) = I$, c'est-à-dire que l'intégrale (8) converge.

Supposons maintenant que l'intégrale (8) diverge. Si l'on suppose que l'intégrale (7) converge, il en sera de même de l'intégrale (8) d'après ce qui précède. Or ceci est contraire à l'hypothèse. Donc l'intégrale (7) diverge aussi. ■

Remarque. On peut formuler de façon analogue un critère de comparaison des intégrales impropres de seconde espèce : si $f(x)$ et $g(x)$ sont des fonctions continues sur un intervalle $]a, b]$ et si $0 \leqslant f(x) \leqslant g(x)$ sur un intervalle $]a, a + \varepsilon[$, alors la convergence de l'intégrale $\int_a^b g(x)\,dx$ entraîne celle de l'intégrale $\int_a^b f(x)\,dx$ et la divergence de l'intégrale $\int_a^b f(x)\,dx$ celle de l'intégrale $\int_a^b g(x)\,dx$.

Exemple 6. Etudier la convergence de l'intégrale $\int_1^{+\infty} \frac{dx}{x^2(1+e^{-x})}$.

Solution. Comparons l'intégrant $\frac{1}{x^2(1+e^{-x})}$ à la fonction $1/x^2$ sur l'intervalle $[1, +\infty[$. Il est évident que

$$\frac{1}{x^2(1+e^{-x})} < \frac{1}{x^2}.$$

Or l'intégrale $\int_1^{+\infty} \frac{dx}{x^2}$ converge, puisque $\alpha = 2 > 1$ (cf. exemple 4). Donc il en est de même de l'intégrale envisagée d'après le critère de comparaison.

Exemple 7. Etudier la convergence de l'intégrale $\int_1^{+\infty} \frac{\sqrt{x}}{1+x}\,dx$.

Solution. En comparant l'intégrant $\frac{\sqrt{x}}{1+x}$ à la fonction $\frac{1}{2\sqrt{x}}$ sur l'intervalle $[1, +\infty[$, on trouve que

$$\frac{1}{2\sqrt{x}} = \frac{\sqrt{x}}{x+x} \leqslant \frac{\sqrt{x}}{1+x}.$$

Or l'intégrale $\int_1^{+\infty} \frac{dx}{2\sqrt{x}}$ diverge, puisque $\alpha = \frac{1}{2} < 1$ (cf. exemple 4). Donc il en est de même de l'intégrale considérée en vertu du critère de comparaison.

4. Exemple d'application de l'intégrale impropre. Calculons la vitesse de libération d'un véhicule spatial.

Nous avons calculé précédemment (cf. § 10, n° 6, exemple 11) à l'aide de l'intégrale définie le travail nécessaire pour lancer un corps de masse m à une altitude h:

$$A = \int_R^{R+h} F(x)\,dx = \frac{PRh}{R+h}.$$

L'envoi d'un engin dans l'espace interplanétaire signifie que $h = \infty$. Calculons le travail correspondant:

$$\lim_{h\to\infty} A = \int_R^{\infty} F(x)\,dx = \lim_{h\to\infty} \frac{PRh}{R+h} = \lim_{h\to\infty} \frac{PR}{1+R/h} = PR = mgR,$$

où m est la masse de l'engin spatial; g, l'accélération de la pesanteur terrestre (le frottement et l'attraction des autres planètes sont négligés). Ce travail est accompli grâce à la variation de l'énergie cinétique de l'engin. Donc cette énergie doit être à l'instant initial supérieure à ce travail, c'est-à-dire que la vitesse de libération v doit être telle que

$$\frac{mv^2}{2} \geqslant mgR \text{ ou } v \geqslant \sqrt{2gR} = \sqrt{2 \cdot 10 \cdot 6\,400\,000} \text{ m/s} =$$

$$= 1{,}4 \cdot 8000 \text{ m/s} = 11{,}2 \text{ km/s}.$$

La trajectoire du véhicule spatial est une parabole, une hyperbole ou une ellipse selon que la vitesse de libération est égale, supérieure ou inférieure à 11,2 km/s. Dans le dernier cas, l'engin peut soit s'écraser sur la Terre, soit en devenir un satellite artificiel.

§ 12. Calcul approché des intégrales définies

La résolution des problèmes de physique et autres implique le calcul d'intégrales définies de fonctions dont les primitives ne s'expriment pas par des fonctions élémentaires. D'où la nécessité d'établir des formules de calcul approché des intégrales définies. Voyons deux d'entre elles: la formule des trapèzes et la formule des paraboles.

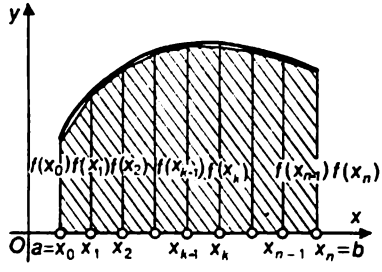


Fig. 118

1. Formule des trapèzes ou de Poncelet. Soit à calculer l'intégrale $\int_a^b f(x)\,dx$, où $f(x)$ est une fonction continue. Pour alléger les raisonnements on se bornera au cas où $f(x) \geqslant 0$. Divisons l'intervalle $[a, b]$ en n parties égales par les points $a = x_0 < x_1 < x_2 < \ldots < x_{k-1} < x_k < \ldots < x_n = b$ et construisons n trapèzes rectangles en traçant les droites $x = x_k$ (ces trapèzes sont hachurés sur la figure 118). La somme des aires des trapèzes est égale approximativement à l'aire du trapèze curviligne, c'est-à-dire que

$$\int_a^b f(x)\,dx \approx \frac{f(x_0)+f(x_1)}{2}(x_1-x_0)+\frac{f(x_1)+f(x_2)}{2}(x_2-x_1)+\ldots$$

$$\ldots+\frac{f(x_{n-1})+f(x_n)}{2}(x_n-x_{n-1})=\frac{b-a}{2n}\left\{f(a)+f(b)+2\sum_{k=1}^{n-1}f(x_k)\right\},$$

où $f(x_{k-1})$ et $f(x_k)$ sont les bases respectives des trapèzes; $x_k - x_{k-1} = (b-a)/n$ leurs hauteurs.

On a ainsi obtenu la formule approchée

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2n}\left\{f(a)+f(b)+2\sum_{k=1}^{n-1}f(x_k)\right\},$$

qui s'appelle *formule des trapèzes* ou *formule de Poncelet* *). Cette formule est d'autant plus exacte que n est grand.

Calculons à titre d'exemple l'intégrale $\int_0^1 x^2\,dx$. La valeur de cette intégrale s'obtient sans peine:

$$\int_0^1 x^2\,dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3} = 0{,}3333\ldots$$

*) Poncelet Jean Victor (1788-1867), mathématicien français.

Trouvons maintenant une valeur approchée à l'aide de la formule des trapèzes. Soit $n = 5$. On a alors $a = x_0 = 0$, $x_1 = 0{,}2$, $x_2 = 0{,}4$, $x_3 = 0{,}6$, $x_4 = 0{,}8$, $x_5 = 1 = b$ et respectivement $f(x_0) = 0$, $f(x_1) = 0{,}04$, $f(x_2) = 0{,}16$, $f(x_3) = 0{,}36$, $f(x_4) = 0{,}64$, $f(x_5) = 1$. Donc

$$\int_0^1 x^2\,dx \approx \frac{1}{10}\{1 + 2[0{,}04 + 0{,}16 + 0{,}36 + 0{,}64]\} = \frac{3{,}4}{10} = 0{,}34.$$

La valeur exacte étant égale à 0,3333 . . ., l'erreur absolue est $<0{,}007$. Cette précision suffit dans de nombreux problèmes pratiques.

La précision croît avec n. Ainsi, par exemple, pour $n = 10$

$$\int_0^1 x^2\,dx \approx \frac{1}{20}\{1 + 2 \cdot 2{,}85\} = \frac{6{,}7}{20} = 0{,}335,$$

c'est-à-dire que l'erreur absolue est $<0{,}002$.

Dans les cours de mathématiques supérieures plus complets on démontre que si la fonction $f(x)$ admet une dérivée seconde continue sur $[a, b]$, l'erreur absolue de la formule des trapèzes est inférieure à

$$k\,\frac{(b-a)^3}{12n^2}$$

où k est la valeur maximale de $|f''(x)|$ sur $[a, b]$.

A signaler que l'accroissement de n entraîne un accroissement non seulement de la précision mais aussi du volume des calculs. Dans ces cas on fait appel aux ordinateurs.

Calculer l'intégrale $\int_0^1 \frac{dx}{1+x}$ à l'aide de la formule des trapèzes pour $n = 10$. Divisons l'intervalle $[0, 1]$ en 10 parties égales par les points $x_0 = 0$, $x_1 = 0{,}1$, . . ., $x_9 = 0{,}9$, $x_{10} = 1$. Calculons approximativement les valeurs prises par la fonction $f(x) = \frac{1}{1+x}$ en ces points: $f(0) = 1{,}0000$, $f(0{,}1) = 0{,}9091$, $f(0{,}2) = 0{,}8333$, $f(0{,}3) = 0{,}7692$, $f(0{,}4) = 0{,}7143$, $f(0{,}5) = 0{,}6667$, $f(0{,}6) = 0{,}6250$, $f(0{,}7) = 0{,}5882$, $f(0{,}8) = 0{,}5556$, $f(0{,}9) = 0{,}5263$, $f(1) = 0{,}5000$.

La formule des trapèzes nous donne

$$\int_0^1 \frac{dx}{1+x} \approx \frac{1}{10}\left(\frac{1{,}000 + 0{,}5000}{2} + 0{,}9091 + 0{,}8333 + \right.$$

$$\left. + 0{,}7692 + 0{,}7143 + 0{,}6667 + 0{,}6250 + 0{,}5882 + 0{,}5556 + 0{,}5263\right) =$$

$$= 0{,}69377 \approx 0{,}6938.$$

Estimons l'erreur du résultat obtenu. On a $f'(x) = -1/(1+x)^2$, $f''(x) = 2/(1+x)^3$. On voit que $|f''(x)| \leqslant 2$ sur $[0, 1]$. Donc l'erreur est inférieure à

$$\frac{k(b-a)^3}{12n^2} = \frac{2}{12 \cdot 10^2} = \frac{1}{600} < 0{,}0017.$$

Calculons maintenant la valeur exacte de cette intégrale à l'aide de la formule de Newton-Leibniz:

$$\int_0^1 \frac{dx}{1+x} = \ln(1+x)\,|_0^1 = \ln 2 \approx 0{,}69315.$$

L'erreur absolue du résultat obtenu par la formule des trapèzes est $<0{,}0007$. Ceci est conforme à l'estimation de l'erreur donnée plus haut.

Le principe de la formule des trapèzes peut être utilisé pour établir des formules plus précises pour le calcul des intégrales définies.

2. Formule de Simpson ou des paraboles. Prouvons préalablement deux lemmes.

Lemme 8.1. *Par trois points quelconques $M_1(x_1; y_1)$, $M_2(x_2; y_2)$ et $M_3(x_3; y_3)$ d'abscisses distinctes il passe une seule courbe de la forme*

$$y = Ax^2 + Bx + C. \tag{1}$$

Démonstration. En portant les coordonnées des points M_1, M_2 et M_3 dans l'équation de la parabole (1), on obtient un système de trois équations du premier degré à trois inconnues A, B et C:

$$\begin{cases} Ax_1^2 + Bx_1 + C = y_1, \\ Ax_2^2 + Bx_2 + C = y_2, \\ Ax_3^2 + Bx_3 + C = y_3. \end{cases}$$

Les nombres x_1, x_2, x_3 étant distincts, le déterminant de ce système (chap. 10, § 3) est non nul:

$$\Delta = \begin{vmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{vmatrix} = (x_2 - x_1)(x_3 - x_2)(x_1 - x_2) \neq 0.$$

Donc ce système admet une seule solution, c'est-à-dire que les coefficients A, B et C sont définis de façon unique. ■

A noter que la courbe (1) est une parabole si $A \neq 0$ et une droite si $A = 0$.

Lemme 8.2. *L'aire s du trapèze curviligne limité par une courbe* $y = Ax^2 + Bx + C$ *passant par les points* $M_1(-h; y_1)$, $M_2(0; y_2)$ *et* $M_3(h; y_3)$ (fig. 119) *est exprimée par la formule*

$$s = \frac{h}{3}(y_1 + 4y_2 + y_3). \tag{2}$$

Démonstration. En portant les coordonnées des points M_1, M_2 et M_3 dans l'équation $y = Ax^2 + Bx + C$, on trouve

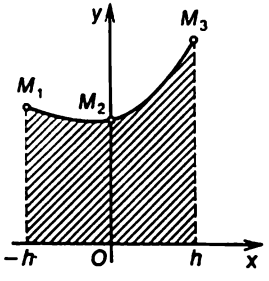


Fig. 119

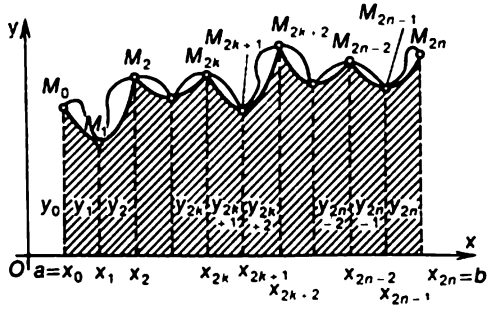


Fig. 120

$y_1 = Ah^2 - Bh + C$; $y_2 = C$; $y_3 = Ah^2 + Bh + C$, d'où

$$2Ah^2 + 2C = y_1 + y_3; \quad C = y_2. \tag{3}$$

En tenant compte des relations (3), on obtient

$$s = \int_{-h}^{h} (Ax^2 + Bx + C)\,dx = \int_{-h}^{h} (Ax^2 + C)\,dx + B\int_{-h}^{h} x\,dx =$$

$$= 2\int_{0}^{h} (Ax^2 + C)\,dx = \frac{h}{3}(2Ah^2 + 6C) = \frac{h}{3}(y_1 + 4y_2 + y_3). \blacksquare$$

Considérons de nouveau le trapèze curviligne limité par une courbe arbitraire $y = f(x)$. Divisons l'intervalle $[a, b]$ en $2n$ segments égaux par les points $a = x_0 < x_1 < x_2 < \ldots < x_{2k} < x_{2k+1} < x_{2k+2} < \ldots < x_{2n-2} < x_{2n-1} < x_{2n} = b$, et la courbe $y = f(x)$, en $2n$ parties par les points M_0, M_1, M_2, $\ldots$, M_{2k}, M_{2k+1}, M_{2k+2}, $\ldots$, M_{2n-2}, M_{2n-1}, M_{2n} (fig. 120) d'intersection des droites $x = x_k$ et de la courbe $y = f(x)$. Par chaque triplet de points

$$M_0M_1M_2, \ \ldots, \ M_{2k}M_{2k+1}M_{2k+2}, \ \ldots, \ M_{2n-2}M_{2n-1}M_{2n}$$

faisons passer une courbe $y = Ax^2 + Bx + C$ (cf. lemme 8.1). On obtient ainsi n trapèzes curvilignes limités par des paraboles ou des droites (ces trapèzes sont hachurés sur la figure 120).

L'aire du trapèze curviligne partiel correspondant à l'intervalle $[x_{2k}, x_{2k+2}]$ étant approximativement égale à celle du trapèze « parabolique » correspondant, la formule (2) nous donne (ici $h = \frac{b-a}{2n}$)

$$\int_{x_{2k}}^{x_{2k+2}} f(x)\,dx \approx \frac{b-a}{6n}(y_{2k} + 4y_{2k+1} + y_{2k+2}),$$

où $y_k = f(x_k)$, $k = 0, 1, 2, \ldots, 2n$. En ajoutant ces égalités terme à terme, on obtient la formule approchée

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6n} \sum_{k=0}^{n-1} (y_{2k} + 4y_{2k+1} + y_{2k+2})$$

ou en détaillant

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6n}[y_0 + y_{2n} + 2(y_2 + y_4 + \ldots + y_{2n-2}) + 4(y_1 + y_3 + \ldots + y_{2n-1})].$$

Cette formule s'appelle *formule de Simpson* *) ou formule *des paraboles.*

La signification géométrique de la formule de Simpson est évidente: l'aire du trapèze curviligne limité par une courbe $y = f(x)$ sur un intervalle $[a, b]$ est approximativement égale à la somme des aires de figures limitées par des paraboles (et des droites).

Dans les cours de mathématiques supérieures plus complets on démontre que si la fonction $f(x)$ admet sur $[a, b]$ une dérivée d'ordre quatre continue, l'erreur absolue de la formule de Simpson est inférieure à

$$M \frac{(b-a)^5}{2880n^4},$$

où M est le maximum de $|f^{(4)}(x)|$ sur $[a, b]$. Nous avons signalé plus haut que l'erreur de la formule des trapèzes était inférieure à

$$k\frac{(b-a)^3}{12n^2} \quad (k = \max_{[a,\,b]} |f''(x)|).$$

Puisque n^4 croît plus vite que n^2, l'erreur de la formule de Simpson diminue bien plus vite que celle de la formule des trapèzes lorsque n croît. Ceci explique que la formule de Simpson fournit une meilleure précision que celle des trapèzes.

*) Simpson Thomas (1710-1761), mathématicien anglais.

Pour comparer la précision des formules d'approximation calculons encore une fois l'intégrale $\int_0^1 \frac{dx}{1+x}$, mais cette fois avec la formule de Simpson pour $n = 4$. Divisons l'intervalle $[0, 1]$ en quatre parties égales par les points $x_0 = 0$, $x_1 = 1/4$, $x_2 = 1/2$, $x_3 = 3/4$, $x_4 = 1$ et calculons approximativement les valeurs de la fonction $f(x) = \frac{1}{1+x}$ en ces points. On obtient $y_0 = 1{,}0000$, $y_1 = 0{,}8000$, $y_2 = 0{,}6662$, $y_3 = 0{,}5714$, $y_4 = 0{,}5000$.

La formule de Simpson nous donne

$$\int_0^1 \frac{dx}{1+x} \approx \frac{b-a}{6n}[y_0+y_1+2y_2+4(y_1+y_3)] =$$

$$= \frac{1-0}{12}[1{,}0000+0{,}5000+2\cdot 0{,}6667+$$

$$+4(0{,}8000+0{,}5714)] \approx 0{,}69325.$$

Estimons l'erreur. La dérivée quatrième de la fonction $f(x) = \frac{1}{1+x}$ est $f^{(4)}(x) = \frac{24}{(1+x)^5}$, d'où il s'ensuit que $|f^{(4)}(x)| \leqslant 24$ sur $[0, 1]$. On peut donc prendre $M = 24$ et l'erreur $24/(2880 \cdot 4^4) < 0{,}0004$. En comparant la valeur exacte à la valeur approchée, on trouve que l'erreur absolue de la formule de Simpson est $< 0{,}00011$. Ceci correspond bien à l'estimation de l'erreur indiquée ci-dessus et montre que la formule de Simpson est sensiblement plus exacte que la formule des trapèzes. Ceci explique l'usage plus fréquent de la formule de Simpson dans le calcul des intégrales définies.

On a signalé plus haut que l'on fait appel aux formules d'approximation pour le calcul des intégrales définies lorsque les primitives ne s'expriment pas par des fonctions élémentaires. Calculons par exemple l'intégrale $\int_0^1 e^{-x^2} dx$ *) par la formule de Simpson à 0,001 près.

Calculons $f^{(4)}(x)$ pour déterminer le nombre $2n$ nécessaire à l'obtention de la précision requise. On a

$$f^{(4)}(x) = 4e^{-x^2}(4x^4 - 12x^2 + 3).$$

D'où l'on déduit que $|f^4(x)| \leqslant 20$ sur $[0, 1]$, puisque $e^{-x^2} \leqslant 1$, $|4x^4 - 12x^2 + 3| \leqslant 5$. On peut donc prendre $M = 20$. La

*) Cette intégrale qui ne s'exprime pas par des fonctions élémentaires joue un rôle important en physique statistique et en théorie de la conductivité calorifique.

formule d'estimation de l'erreur nous donne $20/2880n^4 < 1/1000$, d'où $n^4 > 1000/144$. Pour que cette inégalité soit réalisée, il suffit de prendre $n = 2$, c'est-à-dire que $2n = 4$.

Divisons maintenant l'intervalle [0, 1] en quatre parties égales par les points $x_0 = 0$, $x_1 = 1/4$, $x_2 = 1/2$, $x_3 = 3/4$, $x_4 = 1$ et calculons approximativement les valeurs de la fonction $f(x) = e^{-x^2}$ en ces points: $y_0 = 1{,}0000$, $y_1 = 0{,}9394$, $y_2 = 0{,}7788$, $y_3 = 0{,}5698$, $y_4 = 0{,}3679$. La formule de Simpson nous donne

$$\int_0^1 e^{-x^2}\,dx \approx \frac{1}{12}[1{,}0000 + 0{,}3679 + 2 \cdot 0{,}7788 + 4(0{,}9394 + 0{,}5698)] \approx 0{,}7469.$$

Donc $\int_0^1 e^{-x^2}\,dx \approx 0{,}747$ à 0,001 près. Ainsi en divisant l'intervalle [0, 1] seulement en quatre parties égales et en remplaçant l'intégrale envisagée par la somme du second membre de la formule de Simpson, on obtient la valeur de cette intégrale avec la précision requise.

Signalons en conclusion que chacune de ces méthodes de calcul approché des intégrales renferme un algorithme qui permet d'en faire un large usage sur ordinateur. Ces méthodes sont donc un outil très efficace de calcul des intégrales. S'agissant des intégrales qui ne s'expriment pas par des fonctions élémentaires, on peut les tabuler à l'aide d'un ordinateur et des méthodes élémentaires de calcul approché.

DEUXIÈME PARTIE

ANALYSE MATHÉMATIQUE DES FONCTIONS DE PLUSIEURS VARIABLES

Passons maintenant à l'étude des fonctions de deux et de plus de deux variables. Introduisons préalablement quelques notions de géométrie analytique dans l'espace qui nous serviront à interpréter géométriquement les fonctions de plusieurs variables.

Dans cette partie on envisagera aussi quelques notions d'algèbre supérieure, telles les notions de matrice et de déterminant, qui sont importantes aussi bien en analyse mathématique que dans les autres branches des mathématiques.

CHAPITRE IX

GÉOMÉTRIE ANALYTIQUE DANS L'ESPACE

§ 1. Systèmes de coordonnées rectangulaires dans l'espace

Un système de coordonnées rectangulaires *Oxyz* dans l'espace est défini par la donnée d'une unité de mesure des longueurs et de trois axes *Ox*, *Oy* et *Oz* mutuellement perpendiculaires en un point *O*. Le point *O* s'appelle *origine des coordonnées*, *Ox* l'*axe des abscisses*, *Oy* l'*axe des ordonnées* et *Oz* l'*axe des cotes*.

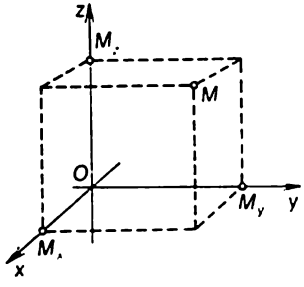


Fig. 121

Soit *M* un point quelconque de l'espace (fig. 121). Menons par *M* trois plans perpendiculaires aux axes de coordonnées *Ox*, *Oy* et *Oz*. Soient M_x, M_y et M_z les points d'intersection de ces plans et des axes. On appelle *coordonnées rectangulaires* du point *M* les nombres

$$x = OM_x \quad y = OM_y, \quad z = OM_z,$$

x est l'*abscisse*, *y* l'*ordonnée*, *z* la *cote* de *M*.

Donc, dans un système de coordonnées donné, à chaque point M de l'espace est associé un triplet ordonné $(x\,;\,y\,;\,z)$ de nombres réels et un seul : ses coordonnées rectangulaires, et réciproquement, à tout triplet ordonné $(x\,;\,y\,;\,z)$ de nombres réels est associé un point M et un seul de l'espace.

Ainsi un système de coordonnées rectangulaires dans l'espace établit une bijection entre l'ensemble des points de l'espace et l'ensemble des triplets ordonnés de nombres.

Les plans *Oxy*, *Oyz* et *Oxz* s'appellent *plans de coordonnées*. Ils partagent l'espace en huit régions appelées *octants*.

§ 2. Notion de vecteur

1. Grandeurs scalaires et vectorielles. De nombreuses grandeurs physiques sont entièrement définies par la donnée d'un seul nombre : le volume, la masse, la densité, la température d'un corps, etc. Ces grandeurs sont appelées *scalaires*. Pour cette raison les nombres sont parfois appelés scalaires. Il existe des grandeurs qui ne peuvent être caractérisées par un simple nombre. Exemple : une force ou une vitesse. De telles grandeurs sont dites *vectorielles*. Pour les décrire on a introduit la notion de vecteur qui est utile en mathématiques.

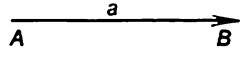


Fig. 122

2. Définitions.

a) *Vecteur lié.* Un couple ordonné de points A et B de l'espace s'appelle *vecteur lié* d'origine A et d'extrémité B (fig. 122).

On le représentera par la notation $\overrightarrow{AB}$. Si les points A et B sont confondus, le vecteur est dit *nul* et se note $\vec{0}$ ou simplement 0.

La distance des points A et B est la *longueur* ou *module* du vecteur $\overrightarrow{AB}$; la droite AB est son *support*. Ce support n'est pas défini pour un vecteur nul.

b) *Vecteurs équipollents.* On dit que deux vecteurs liés sont *équipollents* s'ils ont même direction, même sens et même module.

On établit sans peine que l'équipollence est une relation d'équivalence dans l'ensemble des vecteurs liés.

c) *Vecteur libre.* On appelle *vecteur libre* et on note $\vec{a}$ toute classe d'équivalence de vecteurs liés.

Autrement dit, un vecteur libre est l'ensemble de tous les vecteurs liés équipollents. Il est complètement défini par l'un quelconque de ses représentants.

d) *Vecteurs libres égaux.* On dit que deux vecteurs libres $\vec{a}$ et $\vec{b}$ sont *égaux* si leurs représentants sont équipollents.

La figure 123 représente, à droite, des vecteurs égaux, à gauche, des vecteurs non égaux.

3. Projection d'un vecteur sur un axe. Soient donnés un axe u et un vecteur $\overrightarrow{AB}$. Menons par A et B les plans perpendiculaires à l'axe u et désignons par A' et B' leurs points d'intersection avec u (fig. 124).

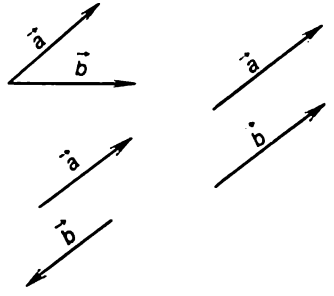


Fig. 123

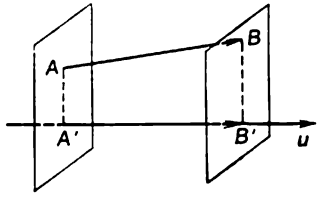


Fig. 124

Le vecteur $\overrightarrow{A'B'}$ est le *projeté* de $\overrightarrow{AB}$ sur l'axe u. On notera ceci: $\overrightarrow{A'B'} = \mathrm{pr}_u \overrightarrow{AB}$. De façon analogue, $\overline{A'B'} = \mathrm{pr}_u \overline{AB}$. On rappelle que

$$\left.\begin{aligned} &\overline{A'B'} = |\overrightarrow{A'B'}| \text{ si } \overrightarrow{A'B'} \text{ et } u \text{ sont de même sens,} \\ &\overline{A'B'} = -|\overrightarrow{A'B'}| \text{ s'ils sont de sens opposé.} \end{aligned}\right\} \quad (1)$$

On a le théorème suivant.

Théorème 9.1. *La mesure algébrique du projeté du vecteur $\overrightarrow{AB}$ sur l'axe u est égale au produit du module du vecteur $\overrightarrow{AB}$ par le cosinus de l'angle du vecteur $\overrightarrow{AB}$ et de l'axe u, c'est-à-dire que*

$$\overline{\mathrm{pr}_u \overrightarrow{AB}} = |\overrightarrow{AB}| \cos\varphi, \quad (2)$$

où $\varphi = (u, \overrightarrow{AB})$ (fig. 125).

Démonstration. Si $\varphi \leqslant \pi/2$ (fig. 125, *a*), alors

$$\overline{\mathrm{pr}_u \overrightarrow{AB}} = |\overrightarrow{A'B'}| = |\overrightarrow{AB}| \cos\varphi.$$

Si $\varphi > \pi/2$ (fig. 125, *b*), on a

$$\overline{\mathrm{pr}_u \overrightarrow{AB}} = -|\overrightarrow{A'B'}| = -|\overrightarrow{AB}| \cos(\pi - \varphi) = |\overrightarrow{AB}| \cos\varphi.$$

La relation (2) est donc vraie pour tout angle φ. ■

Remarque 1. Soit $\overrightarrow{A_1B_1} = \overrightarrow{A_2B_2}$ et soit donné un axe u. En appliquant la formule (2) à chacun de ces vecteurs, on obtient

$$\overline{\mathrm{pr}_u \overrightarrow{A_1B_1}} = \overline{\mathrm{pr}_u \overrightarrow{A_2B_2}},$$

c'est-à-dire que les mesures algébriques des projetés de vecteurs égaux sur un axe sont égales.

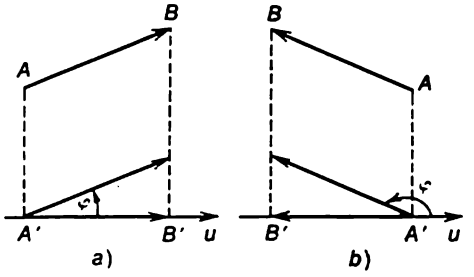


Fig. 125

4. Projection d'un vecteur sur les axes de coordonnées. Soient donnés dans l'espace un système de coordonnées rectangulaires $Oxyz$ et un vecteur $\overrightarrow{AB}$. Soient par ailleurs $X = \overline{\mathrm{pr}_x \overrightarrow{AB}}$, $Y = \overline{\mathrm{pr}_y \overrightarrow{AB}}$, $Z = \overline{\mathrm{pr}_z \overrightarrow{AB}}$. Les nombres X, Y, Z sont appelés *composantes algébriques* du vecteur $\overrightarrow{AB}$. Ce qu'on note

$$\overrightarrow{AB} = \{X; Y; Z\}.$$

Théorème 9.2. *Les composantes d'un vecteur $\overrightarrow{AB}$ d'origine A $(x_1; y_1; z_1)$ et d'extrémité B $(x_2; y_2; z_2)$ s'expriment par les formules*

$$X = x_2 - x_1, \quad Y = y_2 - y_1, \quad Z = z_2 - z_1. \tag{3}$$

Démonstration. Menons par les points A et B des plans perpendiculaires à l'axe Ox et désignons par A' et B' leurs intersections avec Ox. Les abscisses des points A' et B' sont respectivement x_1 et x_2 (fig. 126). Par définition $X = \overline{\mathrm{pr}_x \overrightarrow{AB}} = \overline{A'B'}$. Or $\overline{A'B'} = x_2 - x_1$, donc $X = x_2 - x_1$. On établit de même les autres formules (3). ■

Remarque 2. Si le vecteur $\overrightarrow{AB}$ est appliqué en l'origine des coordonnées, c'est-à-dire que $x_1 = y_1 = z_1 = 0$ et $x_2 = x$, $y_2 = y$ et $z_2 = z$, ses composantes X, Y, Z sont égales aux coordonnées

de son extrémité:

$$X = x, \quad Y = y, \quad Z = z.$$

5. Cosinus directeurs d'un vecteur. Soit donné un vecteur $\vec{a} = \{X; Y, Z\}$ représenté par un vecteur $\overrightarrow{OA}$ dont l'extrémité A n'est située dans aucun plan de coordonnées. Menons par A des plans perpendiculaires aux axes de coordonnées. On obtient ainsi

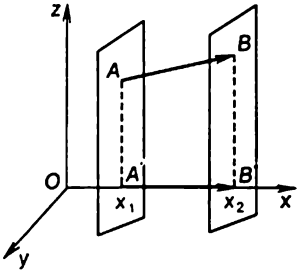


Fig. 126

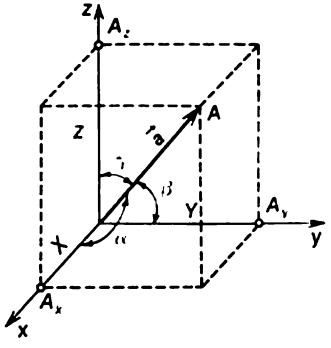


Fig. 127

un parallélépipède de diagonale OA (fig. 127). On sait que le carré de la diagonale d'un parallélépipède rectangle est égal à la somme des carrés de ses arêtes, soit

$$OA^2 = OA_x^2 + OA_y^2 + OA_z^2.$$

Or $OA = |\vec{a}|$, $OA_x = X$, $OA_y = Y$, $OA_z = Z$; donc

$$|\vec{a}|^2 = X^2 + Y^2 + Z^2$$

ou

$$|\vec{a}| = \sqrt{X^2 + Y^2 + Z^2}. \tag{4}$$

La formule (4) exprime la longueur d'un vecteur en fonction de ses composantes.

Désignons par α, β et γ les angles du vecteur $\vec{a}$ et des axes de coordonnées. Les formules (2) et (4) nous donnent

$$\cos\alpha = \frac{X}{\sqrt{X^2+Y^2+Z^2}}, \quad \cos\beta = \frac{Y}{\sqrt{X^2+Y^2+Z^2}},$$

$$\cos\gamma = \frac{Z}{\sqrt{X^2+Y^2+Z^2}}. \tag{5}$$

Ces cosinus s'appellent *cosinus directeurs* du vecteur $\vec{a}$.

En élevant les deux membres de chaque égalité (5) au carré et en sommant les résultats obtenus, on trouve

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1, \tag{6}$$

autrement dit, la somme des carrés des cosinus directeurs de tout vecteur est égale à 1.

En conclusion calculons la distance d de deux points arbitraires $M_1(x_1; y_1; z_1)$ et $M_2(x_2; y_2; z_2)$. Le théorème 9.2 et la formule (4) nous donnent immédiatement le résultat cherché:

$$\overrightarrow{M_1M_2} = \{x_2 - x_1;\ y_2 - y_1;\ z_2 - z_1\},$$

et comme d est la longueur du vecteur $\overrightarrow{M_1M_2}$, on a

$$d = |\overrightarrow{M_1M_2}| = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}. \qquad (7)$$

§ 3. Opérations linéaires sur les vecteurs et leurs propriétés fondamentales

Les opérations linéaires sur les vecteurs sont l'addition et la soustraction des vecteurs ainsi que la multiplication des vecteurs par des scalaires.

1. Addition de deux vecteurs. Soient donnés deux vecteurs libres $\vec{a}$ et $\vec{b}$. On appelle *somme* $\vec{a} + \vec{b}$ des vecteurs $\vec{a}$ et $\vec{b}$ le vecteur libre de représentant $\overrightarrow{OB}$ obtenu en construisant les représentants consécutifs $\overrightarrow{OA}$ de $\vec{a}$ et $\overrightarrow{AB}$ de $\vec{b}$ (fig. 128, *a*).

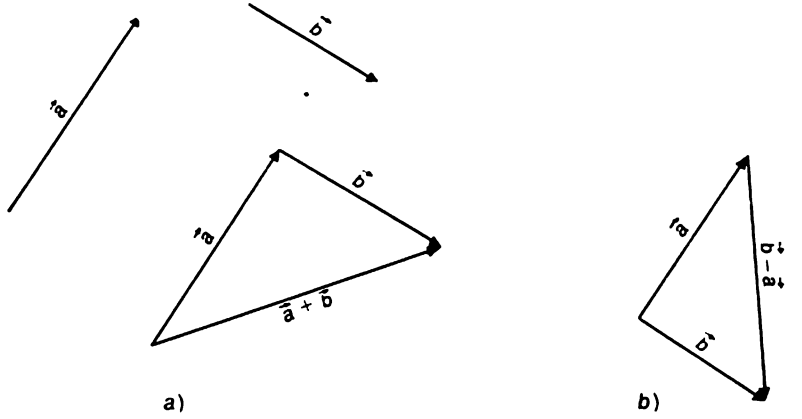


Fig. 128

Remarque 1. La soustraction des vecteurs est l'opération inverse de l'addition, c'est-à-dire que l'on appelle différence $\vec{b} - \vec{a}$ des vecteurs $\vec{b}$ et $\vec{a}$ le vecteur qu'il faut ajouter à $\vec{a}$ pour obtenir $\vec{b}$ (fig. 128, *b*).

Remarque 2. Après avoir défini la somme de deux vecteurs, on peut trouver la somme d'un nombre quelconque de vecteurs. Soient donnés par exemple trois vecteurs $\vec{a}$, $\vec{b}$ et $\vec{c}$. La somme des vecteurs $\vec{a}$ et $\vec{b}$ est le vecteur $\vec{a} + \vec{b}$. En ajoutant le vecteur $\vec{c}$ à $\vec{a} + \vec{b}$ on obtient le vecteur $\vec{a} + \vec{b} + \vec{c}$.

2. Multiplication d'un vecteur par un nombre. Soient donnés un vecteur $\vec{a} \neq \vec{0}$ et un nombre $\lambda \neq 0$. On appelle *produit* $\lambda\vec{a}$ du vecteur $\vec{a}$ par le nombre λ un vecteur de longueur $|\lambda| \cdot |\vec{a}|$, de même direction que $\vec{a}$, de même sens si $\lambda > 0$ et de sens contraire si $\lambda < 0$ (fig. 129).

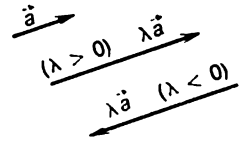


Fig. 129

La multiplication d'un vecteur $\vec{a} \neq \vec{0}$ par un nombre $\lambda \neq 0$ admet l'interprétation géométrique suivante: la multiplication « dilate » le vecteur $\vec{a}$ de λ fois si $|\lambda| > 1$ et le « contracte » de $1/\lambda$ fois si $|\lambda| < 1$. Le vecteur change de sens si $\lambda < 0$. La figure 129 correspond au cas où $|\lambda| > 1$.

On admet que le produit $\lambda\vec{a}$ est nul si l'un des facteurs est nul.

Remarque 3. En se servant de la définition de la multiplication d'un vecteur par un nombre, on démontre sans peine que si deux vecteurs $\vec{a}$ et $\vec{b}$ sont parallèles et si $\vec{a} \neq \vec{0}$, il existe un nombre λ et un seul tel que $\vec{b} = \lambda\vec{a}$ (on laisse cette démonstration au soin du lecteur).

3. Propriétés fondamentales des opérations linéaires.

1° $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ (*commutativité de l'addition*).

Démonstration. Construisons un parallélogramme sur les représentants $\overrightarrow{OA}$ et $\overrightarrow{OB}$ des vecteurs $\vec{a}$ et $\vec{b}$ (fig. 130). On a alors

$$\vec{a} = \overrightarrow{OA}, \quad \vec{b} = \overrightarrow{AC}, \quad \vec{a} + \vec{b} = \overrightarrow{OC},$$

$$\vec{b} = \overrightarrow{OB}, \quad \vec{a} = \overrightarrow{BC}, \quad \vec{b} + \vec{a} = \overrightarrow{OC}.$$

Donc $\overrightarrow{OC} = \vec{a} + \vec{b} = \vec{b} + \vec{a}$. ■

2° $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$ (*associativité de l'addition*).

Démonstration. Construisons les vecteurs $\vec{a} = \overrightarrow{OA}$, $\vec{b} = \overrightarrow{AB}$ et $\vec{c} = \overrightarrow{BC}$ (fig. 131). Alors

$$(\vec{a} + \vec{b}) + \vec{c} = (\overrightarrow{OA} + \overrightarrow{AB}) + \overrightarrow{BC} = \overrightarrow{OB} + \overrightarrow{BC} = \overrightarrow{OC},$$
$$\vec{a} + (\vec{b} + \vec{c}) = \overrightarrow{OA} + (\overrightarrow{AB} + \overrightarrow{BC}) = \overrightarrow{OA} + \overrightarrow{AC} = \overrightarrow{OC}.$$

Donc

$$\overrightarrow{OC} = (\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}).$$

Considérons encore trois autres propriétés dont deux concernent simultanément l'addition des vecteurs et la multiplication

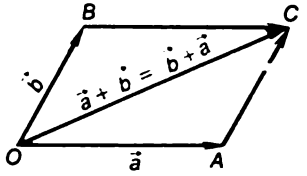


Fig. 130

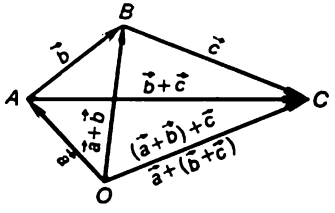


Fig. 131

d'un vecteur par un nombre. Soient λ et μ des nombres, $\vec{a}$ et $\vec{b}$ des vecteurs. Alors

3° $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$ (*associativité de la multiplication*);

4° $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$ (*distributivité par rapport à l'addition des nombres*);

5° $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$ (*distributivité par rapport à l'addition des vecteurs*).

Prouvons la propriété 3°. Les deux membres de l'égalité 3° s'annulent si l'un au moins des facteurs λ, μ et $\vec{a}$ est nul. Si $\lambda \neq 0$, $\mu \neq 0$ et $a \neq 0$, les vecteurs $\lambda(\mu\vec{a})$ et $(\lambda\mu)\vec{a}$ ont la même direction, le même sens (le sens de $\vec{a}$ si λ et μ sont de même signe et le sens contraire de $\vec{a}$ sinon) et la même longueur ($|\lambda(\mu\vec{a})| = |\lambda||\mu\vec{a}| = |\lambda||\mu||\vec{a}|$ et $|(\lambda\mu)\vec{a}| = |\lambda\mu||\vec{a}| = |\lambda||\mu||\vec{a}|$). Donc ils sont égaux. ■

Prouvons la propriété 4°. Supposons que λ et μ sont de même signe et $\vec{a} \neq 0$. Les vecteurs $(\lambda + \mu)\vec{a}$ et $\lambda\vec{a} + \mu\vec{a}$ sont de même direction, de même sens (le sens de $\vec{a}$ pour $\lambda > 0$, $\mu > 0$ et le sens contraire de $\vec{a}$ pour $\lambda < 0$, $\mu < 0$). Les vecteurs $\lambda\vec{a}$ et $\mu\vec{a}$ étant de même sens, la longueur du vecteur $\lambda\vec{a} + \mu\vec{a}$ est égale à

$|\lambda\vec{a} + \mu\vec{a}| = |\lambda\vec{a}| + |\mu\vec{a}| = |\lambda|\,|\vec{a}| + |\mu|\,|\vec{a}| = (|\lambda| + |\mu|)\,\vec{a}$ et puisque λ et μ sont de même signe, il vient $|(\lambda + \mu)\,\vec{a}| = |\lambda + \mu|\,|\vec{a}| = (|\lambda| + |\mu|)\,\vec{a}$, c'est-à-dire que les vecteurs $\lambda\vec{a} + \mu\vec{a}$ et $(\lambda + \mu)\,\vec{a}$ ont même longueur. Donc ils sont égaux.

Supposons maintenant que λ et μ sont de signes contraires et pour fixer les idées que $|\lambda| > |\mu|$. Dans ce cas les vecteurs $(\lambda + \mu)\,\vec{a}$ et $\lambda\vec{a} + \mu\vec{a}$ sont de même sens que le vecteur $\lambda\vec{a}$. La longueur du vecteur $\lambda\vec{a} + \mu\vec{a}$ est égale à $|\lambda\vec{a} + \mu\vec{a}| = |\lambda\vec{a}| - |\mu\vec{a}| = |\lambda|\,|\vec{a}| - |\mu|\,|\vec{a}| = (|\lambda| - |\mu|)\,\vec{a}$, celle du vecteur $(\lambda + \mu)\,\vec{a}$, à $|(\lambda + \mu)\,\vec{a}| = |\lambda + \mu|\,|\vec{a}| = (|\lambda| - |\mu|)\,|\vec{a}|$. Les vecteurs $\lambda\vec{a} + \mu\vec{a}$ et $(\lambda + \mu)\,\vec{a}$ sont donc égaux dans ce cas aussi. Si $|\lambda| = |\mu|$ et si λ et μ sont de signes contraires, les deux membres de l'égalité à prouver sont nuls. L'égalité 4° est évidente si l'un au moins des nombres λ et μ ou le vecteur $\vec{a}$ est nul. ■

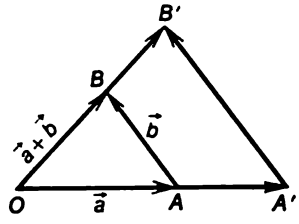


Fig. 132

Prouvons la propriété 5°. Supposons que les vecteurs $\vec{a}$ et $\vec{b}$ ne sont pas parallèles et que $\lambda > 0$. Construisons les vecteurs $\overrightarrow{OB} = \vec{a} + \vec{b}$, $\overrightarrow{OA'} = \lambda\vec{a}$ et $\overrightarrow{OB'} = \lambda(\vec{a} + \vec{b})$ (fig. 132). De la similitude des triangles OAB et $OA'B'$ et de la définition de la multiplication d'un vecteur par un nombre il résulte que $\overrightarrow{A'B'} = \lambda\vec{b}$, et du triangle $OA'B'$, que $\overrightarrow{OA'} + \overrightarrow{A'B'} = \overrightarrow{OB'}$. Donc $\lambda\vec{a} + \lambda\vec{b} = \lambda(\vec{a} + \vec{b})$. Le cas $\lambda < 0$ se traite de façon analogue.

Si $\vec{a}$ et $\vec{b}$ sont parallèles et $\vec{a} \neq \vec{0}$, le vecteur $\vec{b}$ peut être mis sous la forme $\vec{b} = \mu\vec{a}$ et l'égalité cherchée résulte des égalités 3° et 4°. En effet, $\lambda(\vec{a} + \vec{b}) = \lambda(\vec{a} + \mu\vec{a}) = \lambda(1 + \mu)\,\vec{a} = (\lambda + \lambda\mu)\,\vec{a} = \lambda\vec{a} + \lambda\mu\vec{a} = \lambda\vec{a} + \lambda(\mu\vec{a}) = \lambda\vec{a} + \lambda\vec{b}$. La relation à prouver est évidente si l'un des vecteurs $\vec{a}$ et $\vec{b}$ ou si le nombre λ est nul. ■

Remarque 4. Les propriétés prouvées sont fondamentales car elles permettent d'effectuer les opérations algébriques usuelles sur les vecteurs. Par exemple, grâce aux propriétés 4° et 5° on peut multiplier terme à terme un polynôme scalaire par un polynôme vectoriel.

§ 4. Théorèmes des projections

Théorème 9.3. *La mesure algébrique de la projetée de la somme de deux vecteurs sur un axe est égale à la somme des mesures algébriques de leurs projetés sur cet axe, c'est-à-dire que*

$$\overline{\mathrm{pr}_u(\vec{a}_1+\vec{a}_2)} = \overline{\mathrm{pr}_u\,\vec{a}_1} + \overline{\mathrm{pr}_u\,\vec{a}_2}.$$

Démonstration. Soient $\vec{a}_1$ et $\vec{a}_2$ deux vecteurs libres. Construisons $\vec{a}_2 = \overrightarrow{AA_1}$ et $\vec{a}_2 = \overrightarrow{A_1A_2}$ (fig. 133). Désignons par A',

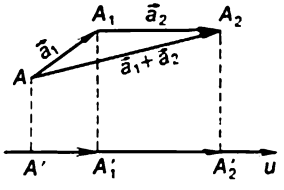


Fig. 133

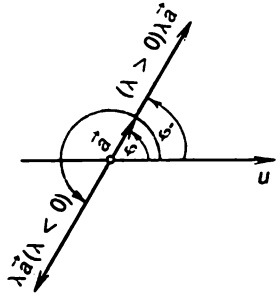


Fig. 134

A'_1 et A'_2 les projetés respectifs des points A, A_1 et A_2 sur l'axe u. Par définition, $\overline{\mathrm{pr}_u\,\vec{a}_1} = \overline{A'A'_1}$, $\overline{\mathrm{pr}_u\,\vec{a}_2} = \overline{A'_1A'_2}$, $\overline{\mathrm{pr}_u(\vec{a}_1+\vec{a}_2)} = \overline{\mathrm{pr}_u\,\overrightarrow{AA_2}} = \overline{A'A'_2}$. D'après la relation de Chasles, $\overline{A'A'_2} = \overline{A'A'_1} + \overline{A'_1A'_2}$. D'où $\overline{\mathrm{pr}_u(\vec{a}_1+\vec{a}_2)} = \overline{\mathrm{pr}_u\,\vec{a}_1} + \overline{\mathrm{pr}_u\,\vec{a}_2}$. ■

Ce théorème peut être généralisé à un nombre quelconque de termes.

Théorème 9.4. *Si un vecteur $\vec{a}$ est multiplié par un nombre $\lambda \neq 0$, la mesure algébrique de son projeté sur un axe u est aussi multipliée par ce nombre, c'est-à-dire que*

$$\overline{\mathrm{pr}_u\,\lambda\vec{a}} = \lambda\,\overline{\mathrm{pr}_u\,\vec{a}}. \qquad (1)$$

Démonstration. Soient $\varphi = (u, \vec{a})$ et $\varphi' = (u, \lambda\vec{a})$ (fig. 134). Si $\lambda > 0$, les vecteurs $\vec{a}$ et $\lambda\vec{a}$ sont de même sens et $\varphi = \varphi'$. Si $\lambda < 0$, les vecteurs $\vec{a}$ et $\lambda\vec{a}$ sont de sens contraire et $\varphi' = \pi + \varphi$. D'après le théorème 9.1, on a :

pour $\lambda > 0$

$$\overline{\mathrm{pr}_u\,\lambda\vec{a}} = |\,\lambda\vec{a}\,|\cos\varphi' = \lambda\,|\vec{a}|\cos\varphi' = \lambda\,|\vec{a}|\cos\varphi = \lambda\,\overline{\mathrm{pr}_u\,\vec{a}};$$

pour $\lambda < 0$

$$\overline{\mathrm{pr}_u\, \lambda\vec{a}} = |\lambda\vec{a}| \cos\varphi' = |\lambda|\,|\vec{a}|\cos\varphi' = -\lambda\,|\vec{a}|\cos(\pi+\varphi) =$$

$$= -\lambda\,|\vec{a}|\,(-\cos\varphi) = \lambda\,|\vec{a}|\cos\varphi = \lambda\,\overline{\mathrm{pr}_u\,\vec{a}};$$

l'égalité (1) est évidente pour $\lambda = 0$. Donc, $\overline{\mathrm{pr}_u\,\lambda\vec{a}} = \lambda\,\overline{\mathrm{pr}_u\,\vec{a}}$ quel que soit $\lambda \neq 0$. ■

Les théorèmes prouvés entraînent deux importants corollaires.

Corollaire 1. *Du théorème* 9.3 *il s'ensuit que si* $\vec{a} = \{X_1; Y_1; Z_1\}$ *et* $\vec{b} = \{X_2; Y_2; Z_2\}$, *alors* $\vec{a} + \vec{b} = \{X_1 + X_2; Y_1 + Y_2; Z_1 + Z_2\}$.

Corollaire 2. *Du théorème* 9.4 *il résulte que si* $\vec{a} = \{X; Y; Z\}$, *alors* $\lambda\vec{a} = \{\lambda X; \lambda Y; \lambda Z\}$ *pour tout nombre* λ.

De là on déduit sans peine une condition de parallélisme (ou de colinéarité) de deux vecteurs. En effet, l'égalité $\vec{b} = \lambda\vec{a}$ équivaut aux égalités $X_2 = \lambda X_1$, $Y_2 = \lambda Y_1$, $Z_2 = \lambda Z_1$ ou

$$\frac{X_2}{X_1} = \frac{Y_2}{Y_1} = \frac{Z_2}{Z_1}, \qquad (2)$$

autrement dit les vecteurs $\vec{a}$ et $\vec{b}$ sont parallèles (colinéaires) si et seulement si leurs composantes sont proportionnelles.

§ 5. Décomposition d'un vecteur suivant une base

Soient $\vec{i}$, $\vec{j}$ et $\vec{k}$ les vecteurs unitaires des axes de coordonnées, c'est-à-dire $|\vec{i}| = |\vec{j}| = |\vec{k}| = 1$ (fig. 135). Le triplet de vecteurs $\vec{i}$, $\vec{j}$, $\vec{k}$ s'appelle *base*. On a le théorème suivant.

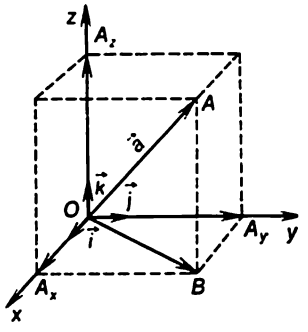


Fig. 135

Théorème 9.5. *Tout vecteur* $\vec{a}$ *peut être décomposé d'une seule manière suivant la base* $\vec{i}$, $\vec{j}$, $\vec{k}$, *c'est-à-dire représenté sous la forme*

$$\vec{a} = \lambda\vec{i} + \mu\vec{j} + \nu\vec{k}, \qquad (1)$$

où λ, μ *et* ν *sont des nombres algébriques.*

Démonstration. Soit $\vec{a}$ un vecteur libre. Construisons le vecteur $\vec{a} = \overrightarrow{OA}$. Menons par A les plans perpendiculaires aux axes de coordonnées et désignons par A_x, A_y et A_z leurs points

d'intersection respectifs avec les axes de coordonnées. Par définition de l'addition des vecteurs, on a

$$\vec{a} = \overrightarrow{OB} + \overrightarrow{OA}_z, \quad \overrightarrow{OB} = \overrightarrow{OA}_x + \overrightarrow{OA}_y. \tag{2}$$

D'où

$$\vec{a} = \overrightarrow{OA}_x + \overrightarrow{OA}_y + \overrightarrow{OA}_z. \tag{3}$$

Les vecteurs $\overrightarrow{OA}_x$ et $\vec{i}$, $\overrightarrow{OA}_y$ et $\vec{j}$, $\overrightarrow{OA}_z$ et $\vec{k}$ étant colinéaires, on a

$$\overrightarrow{OA}_x = \lambda\vec{i}, \quad \overrightarrow{OA}_y = \mu\vec{j}, \quad \overrightarrow{OA}_z = \nu\vec{k}, \tag{4}$$

où λ, μ et ν sont des nombres algébriques. Les relations (3) et (4) entraînent

$$\vec{a} = \lambda\vec{i} + \mu\vec{j} + \nu\vec{k}.$$

Pour prouver l'unicité de la représentation (1), montrons que

$$\lambda = X, \quad \mu = Y, \quad \nu = Z,$$

où X, Y et Z sont les composantes du vecteur $\vec{a}$.

Montrons par exemple que $\lambda = X$. Puisque $X = OA_x$ si $\overrightarrow{OA}_x$ est de même sens que $\vec{i}$ et $X = -OA_x$ s'il est de sens contraire, il vient $\overrightarrow{OA}_x = X\vec{i}$. En comparant avec l'égalité $\overrightarrow{OA}_x = \lambda\vec{i}$, on trouve que $\lambda = X$.

On démontre de façon analogue que $\mu = Y$ et $\nu = Z$. ■

§ 6. Produit scalaire des vecteurs

1. Définition et propriétés fondamentales du produit scalaire.

Définition. On appelle *produit scalaire* de deux vecteurs non nuls $\vec{a}$ et $\vec{b}$ le nombre égal au produit des longueurs de ces vecteurs par le cosinus de l'angle qu'ils forment. Si l'un au moins des vecteurs est nul, l'angle n'est pas défini et le produit scalaire est par définition nul.

Le produit scalaire des vecteurs $\vec{a}$ et $\vec{b}$ se note $\vec{a}\cdot\vec{b}$. Ainsi

$$\vec{a}\cdot\vec{b} = |\vec{a}|\,|\vec{b}|\cos\varphi,$$

où φ est l'angle des vecteurs $\vec{a}$ et $\vec{b}$ (fig. 136).

Puisque $|\vec{a}|\cos\varphi = \overline{\mathrm{pr}_{\vec{b}}\,\vec{a}}$, et $|\vec{b}|\cos\varphi = \overline{\mathrm{pr}_{\vec{a}}\,\vec{b}}$, on peut écrire

$$\vec{a}\cdot\vec{b} = |\vec{b}|\,\overline{\mathrm{pr}_{\vec{b}}\,\vec{a}} = |\vec{a}|\,\overline{\mathrm{pr}_{\vec{a}}\,\vec{b}}. \tag{1}$$

Un exemple classique de produit scalaire en physique est la formule du travail

$$A = |\vec{a}|\,|\vec{b}|\cos\varphi,$$

où $\vec{a}$ est un vecteur force dont le point d'application se déplace de l'origine à l'extrémité d'un représentant de $\vec{b}$ (fig. 137).

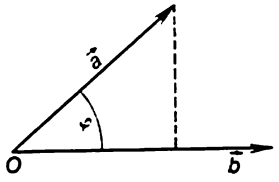


Fig. 136

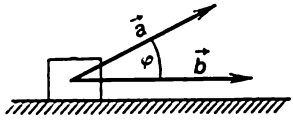


Fig. 137

Considérons quelques propriétés du produit scalaire.

1° $\vec{a}\cdot\vec{b} = \vec{b}\cdot\vec{a}$ (*commutativité*).

Démonstration. Par définition du produit scalaire, $\vec{a}\cdot\vec{b} = |\vec{a}|\,|\vec{b}|\cos\varphi$ et $\vec{b}\cdot\vec{a} = |\vec{b}|\,|\vec{a}|\cos(-\varphi)$, mais $|\vec{a}|\,|\vec{b}| = |\vec{b}|\,|\vec{a}|$ puisque c'est un produit de nombres et $\cos\varphi = \cos(-\varphi)$. Donc $\vec{a}\cdot\vec{b} = \vec{b}\cdot\vec{a}$. ■

2° $(\lambda a)\cdot\vec{b} = \lambda(\vec{a}\cdot\vec{b})$ (*associativité par rapport au produit par un nombre*).

Démonstration. La formule (1) nous donne

$$(\lambda\vec{a})\cdot\vec{b} = |\vec{b}|\,\overline{\mathrm{pr}_{\vec{b}}(\lambda\vec{a})},$$

mais d'après le théorème 9.4

$$\overline{\mathrm{pr}_{\vec{b}}(\lambda\vec{a})} = \lambda\,\overline{\mathrm{pr}_{\vec{b}}\vec{a}}.$$

Donc

$$(\lambda\vec{a})\cdot\vec{b} = |\vec{b}|\,\overline{\mathrm{pr}_{\vec{b}}(\lambda\vec{a})} = |\vec{b}|\,\lambda\,\overline{\mathrm{pr}_{\vec{b}}\vec{a}} = \lambda\,(|\vec{b}|\,\overline{\mathrm{pr}_{\vec{b}}\vec{a}}).$$

Par ailleurs, la même formule (1) nous donne

$$|\vec{b}|\,\overline{\mathrm{pr}_{\vec{b}}\vec{a}} = \vec{a}\cdot\vec{b}.$$

Donc $(\lambda\vec{a})\cdot\vec{b} = \lambda(|\vec{b}|\,\overline{\mathrm{pr}_{\vec{b}}\vec{a}}) = \lambda(\vec{a}\cdot\vec{b})$. ■

Remarque 1. Les propriétés 1° et 2° entraînent $(\lambda\vec{a})\cdot(\mu\vec{b}) = (\lambda\cdot\mu)(\vec{a}\cdot\vec{b})$. En effet, $(\lambda\vec{a})\cdot(\mu\vec{b}) = \lambda[\vec{a}\cdot(\mu\vec{b})] = \lambda[(\mu\vec{b})\cdot\vec{a}] = \lambda[\mu(\vec{b}\cdot\vec{a})] = (\lambda\cdot\mu)(\vec{a}\cdot\vec{b})$.

3° $\vec{a}\cdot(\vec{b}+\vec{c}) = \vec{a}\cdot\vec{b} + \vec{a}\cdot\vec{c}$ (*distributivité par rapport à la somme des vecteurs*).

Démonstration. D'après la formule (1)

$$\vec{a}\cdot(\vec{b}+\vec{c}) = |\vec{a}|\,\overline{\mathrm{pr}_{\vec{a}}(\vec{b}+\vec{c})},$$

d'après le théorème 9.3

$$\overline{\mathrm{pr}_{\vec{a}}(\vec{b}+\vec{c})} = \overline{\mathrm{pr}_{\vec{a}}\vec{b}} + \overline{\mathrm{pr}_{\vec{a}}\vec{c}}.$$

Donc

$$\vec{a}\cdot(\vec{b}+\vec{c}) = |\vec{a}|\,\overline{\mathrm{pr}_{\vec{a}}(\vec{b}+\vec{c})} = |\vec{a}|\,(\overline{\mathrm{pr}_{\vec{a}}\vec{b}} + \overline{\mathrm{pr}_{\vec{a}}\vec{c}}) =$$

$$= |\vec{a}|\,\overline{\mathrm{pr}_{\vec{a}}\vec{b}} + |\vec{a}|\,\overline{\mathrm{pr}_{\vec{a}}\vec{c}}.$$

La même formule (1) nous donne

$$|\vec{a}|\,\overline{\mathrm{pr}_{\vec{a}}\vec{b}} = \vec{a}\cdot\vec{b} \quad \text{et} \quad |\vec{a}|\,\overline{\mathrm{pr}_{\vec{a}}\vec{c}} = \vec{a}\cdot\vec{c}.$$

Donc

$$\vec{a}\cdot(\vec{b}+\vec{c}) = |\vec{a}|\,\overline{\mathrm{pr}_{\vec{a}}\vec{b}} + |\vec{a}|\,\overline{\mathrm{pr}_{\vec{a}}\vec{c}} = \vec{a}\cdot\vec{b} + \vec{a}\cdot\vec{c}. \quad \blacksquare$$

Remarque 2. Cette propriété permet de multiplier les polynômes vectoriels terme à terme. D'après la propriété 1° on peut ne pas se soucier de l'ordre des facteurs et d'après la propriété 2° cf. remarque 1) on peut regrouper les coefficients numériques des facteurs. Exemple:

$$(2\vec{a}+5\vec{b})\cdot(3\vec{c}+4\vec{d}) = (2\vec{a}+5\vec{b})\cdot(3\vec{c}) + (2\vec{a}+5\vec{b})\cdot(4\vec{d}) =$$

$$= (2\vec{a})\cdot(3\vec{c}) + (5\vec{b})\cdot(3\vec{c}) + (2\vec{a})\cdot(4\vec{d}) + (5\vec{b})\cdot(4\vec{d}) =$$

$$= 6\vec{a}\cdot\vec{c} + 15\vec{b}\cdot\vec{c} + 8\vec{a}\cdot\vec{d} + 20\vec{b}\cdot\vec{d}.$$

4°. $\vec{a}\cdot\vec{a} = |\vec{a}|^2$.

Démonstration. Par définition du produit scalaire, $\vec{a}\cdot\vec{a} = |\vec{a}|\cdot|\vec{a}|\cos 0 = |\vec{a}|^2$ si $|\vec{a}| \neq 0$, c'est-à-dire si $\vec{a} \neq \vec{0}$. Si $\vec{a} = \vec{0}$, on a de même par définition $\vec{a}\cdot\vec{a} = 0$. Mais dans ce cas $|\vec{a}| = 0$ et par suite l'égalité $\vec{a}\cdot\vec{a} = |\vec{a}|^2$ a lieu aussi. $\blacksquare$

Le produit scalaire $\vec{a}\cdot\vec{a}$ s'appelle *carré scalaire* du vecteur $\vec{a}$ et se note $\vec{a}^2$. D'après ce qui précède: $\vec{a}^2 = |\vec{a}|^2$, d'où en particulier $\sqrt{\vec{a}^2} = |\vec{a}|$.

5° $\vec{a}\cdot\vec{b}=0$ si $\vec{a}\perp\vec{b}$ et réciproquement, $\vec{a}\perp\vec{b}$ si $\vec{a}\cdot\vec{b}=0$, $\vec{a}\neq 0$, $\vec{b}\neq 0$.

Démonstration. Par définition du produit scalaire, $\vec{a}\cdot\vec{b}=|\vec{a}|\cdot|\vec{b}|\cos\varphi$. Si $\varphi=\pi/2$, c'est-à-dire si les vecteurs $\vec{a}$ et $\vec{b}$ sont orthogonaux, alors $\cos\varphi=0$, d'où $\vec{a}\cdot\vec{b}=0$.

Réciproquement, si $\vec{a}\cdot\vec{b}=0$ et $|\vec{a}|\,|\vec{b}|\neq 0$, alors $\cos\varphi=0$ et $\varphi=\pi/2$, donc les vecteurs $\vec{a}$ et $\vec{b}$ sont orthogonaux. ■

Remarque 3. En appliquant les propriétés 4° et 5° aux vecteurs de base $\vec{i}$, $\vec{j}$ et $\vec{k}$ (fig. 138), on trouve immédiatement:

$$\vec{i}^2=\vec{j}^2=\vec{k}^2=1,$$
$$\vec{i}\cdot\vec{j}=\vec{i}\cdot\vec{k}=\vec{j}\cdot\vec{i}=\vec{j}\cdot\vec{k}=$$
$$=\vec{k}\cdot\vec{i}=\vec{k}\cdot\vec{j}=0. \qquad (2)$$

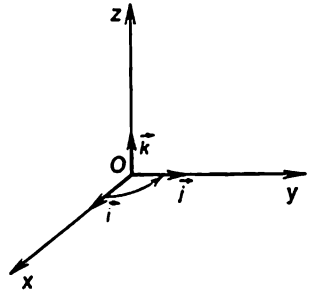


Fig. 138

2. Expression analytique du produit scalaire.

Théorème 9.6. *Si des vecteurs $\vec{a}$ et $\vec{b}$ sont définis par leurs composantes $\vec{a}=\{X_1;\ Y_1;\ Z_1\}$ et $\vec{b}=\{X_2;\ Y_2;\ Z_2\}$, leur produit scalaire s'exprime par la formule*

$$\vec{a}\cdot\vec{b}=X_1X_2+Y_1Y_2+Z_1Z_2.$$

Démonstration. Décomposons les vecteurs $\vec{a}$ et $\vec{b}$ suivant la base $\vec{i}$, $\vec{j}$, $\vec{k}$:

$$\vec{a}=X_1\vec{i}+Y_1\vec{j}+Z_1\vec{k},$$
$$\vec{b}=X_2\vec{i}+Y_2\vec{j}+Z_2\vec{k}.$$

La remarque 2 nous donne

$$\vec{a}\cdot\vec{b}=X_1X_2\vec{i}^2+X_1Y_2\vec{i}\cdot\vec{j}+X_1Z_2\vec{i}\cdot\vec{k}+Y_1X_2\vec{j}\cdot\vec{i}+Y_1Y_2\vec{j}^2+$$
$$+Y_1Z_2\vec{j}\cdot\vec{k}+Z_1X_2\vec{k}\cdot\vec{i}+Z_1Y_2\vec{k}\cdot\vec{j}+Z_1Z_2\vec{k}^2.$$

D'où grâce à (2)

$$\vec{a}\cdot\vec{b}=X_1X_2+Y_1Y_2+Z_1Z_2. \quad ■$$

Le théorème 9.6 entraîne les deux importants corollaires suivants.

Corollaire 1. *Pour que deux vecteurs* $\vec{a} = \{X_1; Y_1; Z_1\}$ *et* $\vec{b} = \{X_2; Y_2; Z_2\}$ *soient orthogonaux, il est nécessaire et suffisant que*

$$X_1X_2 + Y_1Y_2 + Z_1Z_2 = 0. \tag{3}$$

Cette proposition résulte directement de la propriété 5° et du théorème 9.6.

Corollaire 2. *L'angle des vecteurs* $\vec{a} = \{X_1; Y_1; Z_1\}$ *et* $\vec{b} = \{X_2; Y_2; Z_2\}$ *s'exprime par la formule*

$$\cos\varphi = \frac{X_1X_2+Y_1Y_2+Z_1Z_2}{\sqrt{X_1^2+Y_1^2+Z_1^2}\,\sqrt{X_2^2+Y_2^2+Z_2^2}}. \tag{4}$$

En effet, par définition du produit scalaire, $\vec{a}\cdot\vec{b} = |\vec{a}|\,|\vec{b}|\cos\varphi$, d'où

$$\cos\varphi = \frac{\vec{a}\cdot\vec{b}}{|\vec{a}|\,|\vec{b}|}. \tag{5}$$

D'où l'on déduit la formule (4) en vertu du théorème 9.6 et de la formule (4) du § 2.

Exemple. Soient donnés les points A (1; 1; 1), B (2; 2; 1) e C (2; 1; 2). Calculer l'angle $\varphi = \widehat{BAC}$.

Solution. Le théorème 9.2 nous donne $\vec{AB} = \{1; 1; 0\}$, $\vec{AC} = \{1; 0; 1\}$. D'où en vertu de la formule (4)

$$\cos\varphi = \frac{1\cdot1+1\cdot0+0\cdot1}{\sqrt{1^2+1^2+0^2}\,\sqrt{1^2+0^2+1^2}} = \frac{1}{\sqrt{2}\sqrt{2}} = \frac{1}{2}.$$

Donc $\varphi = \pi/3$.

§ 7. Produit vectoriel

1. Définition du produit vectoriel. On dit que des vecteurs libres $\vec{a}$, $\vec{b}$ et $\vec{c}$ sont *coplanaires* s'ils sont parallèles à un même plan.

On dit qu'un trièdre $\vec{i}, \vec{j}, \vec{k}$ *effectif*, c'est-à-dire formé de vecteurs non coplanaires, est *direct* si la rotation, d'un angle inférieur à π, autour de $\vec{k}$, qui amène $\vec{i}$ à coïncider avec $\vec{j}$ s'effectue de la droite vers la gauche pour un observateur placé suivant $\vec{k}$. Le trièdre est *rétrograde* dans le cas contraire.

Définition. On appelle *produit vectoriel* de deux vecteurs libres $\vec{a}$ et $\vec{b}$ un vecteur $\vec{a} \wedge \vec{b}$ dont

1) le module est égal à $|\vec{a}|\,|\vec{b}|\sin\varphi$, où φ est l'angle des vecteurs $\vec{a}$ et $\vec{b}$;

2) la direction est orthogonale à $\vec{a}$ et $\vec{b}$;

3) le sens est tel que le trièdre $\vec{a}$, $\vec{b}$, $\vec{a}\wedge\vec{b}$ est direct (fig. 139).

Signalons que les conditions 2) et 3) se rapportent au cas où $|\vec{a}|\,|\vec{b}|\sin\varphi \neq 0$. Si $|\vec{a}|\,|\vec{b}|\sin\varphi = 0$ (c'est-à-dire que ou

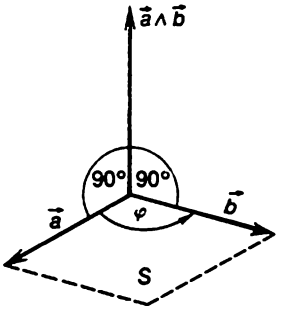


Fig. 139

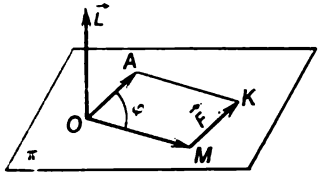


Fig. 140

bien l'un des vecteurs $\vec{a}$ et $\vec{b}$ est nul, ou bien $\sin\varphi = 0$), le produit vectoriel $\vec{a}\wedge\vec{b}$ est défini par la seule condition 1): dans ce cas $\vec{a}\wedge\vec{b} = \vec{0}$.

La notion de produit vectoriel a été empruntée à la mécanique.

Supposons qu'une force $\vec{F} = \overrightarrow{MK}$ est appliquée en un point M d'un solide et soit O un point de l'espace. On sait que le moment de la force $\vec{F}$ par rapport à O est par définition le vecteur $\vec{L}$ (appliqué en O) dont

1) le module est égal à $|\overrightarrow{OM}|\,|\overrightarrow{MK}|\sin\varphi$, où φ est l'angle des vecteurs $\overrightarrow{OM}$ et $\overrightarrow{MK}$;

2) la direction est perpendiculaire au plan π passant par les points O, M, K;

3) le sens est tel que le trièdre $\overrightarrow{OM}$, $\overrightarrow{OA}$, $\overrightarrow{OL}$ est direct (fig. 140).

On voit sur la figure que $\vec{L} = \overrightarrow{OM}\wedge\overrightarrow{MK}$, où $\overrightarrow{OA} = \overrightarrow{MK}$.

2. Propriétés fondamentales du produit vectoriel.

1° $\vec{a}\wedge\vec{b} = \vec{0}$ si les vecteurs $\vec{a}$ et $\vec{b}$ sont colinéaires.

Démonstration. Si les vecteurs $\vec{a}$ et $\vec{b}$ sont parallèles, on a $\sin\varphi = 0$. Donc $|\vec{a}\wedge\vec{b}| = |\vec{a}|\,|\vec{b}|\sin\varphi = 0$. Ce qui exprime que $\vec{a}\wedge\vec{b} = \vec{0}$. ■

2° Le module du produit vectoriel de deux vecteurs $\vec{a}$ et $\vec{b}$ non parallèles est égal à l'aire s du parallélogramme construit sur ces vecteurs (fig. 139).

Démonstration. On sait que l'aire d'un parallélogramme est égale au produit de deux côtés adjacents par le sinus de leur angle. D'où $|\vec{a}|\,|\vec{b}|\sin\varphi = s$, c'est-à-dire que $|\vec{a} \wedge \vec{b}| = s$. ■

3° $\vec{a} \wedge \vec{b} = -\vec{b} \wedge \vec{a}$ (*anticommutativité*).

Démonstration. Cette propriété est évidente si les vecteurs $\vec{a}$ et $\vec{b}$ sont colinéaires. S'ils ne le sont pas, il s'ensuit de la définition du produit vectoriel que les vecteurs $\vec{a} \wedge \vec{b}$ et $\vec{b} \wedge \vec{a}$

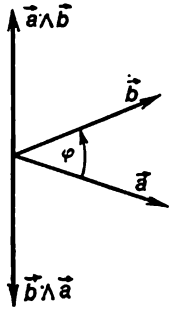


Fig. 141

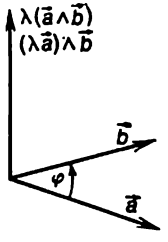


Fig. 142

ont même module (le module du produit vectoriel est indépendant de l'ordre des facteurs), sont parallèles (car orthogonaux à un même plan), mais de sens contraires (fig. 141), puisque les trièdres $\vec{a}$, $\vec{b}$, $\vec{a} \wedge \vec{b}$ et $\vec{b}$, $\vec{a}$, $\vec{b} \wedge \vec{a}$ sont directs. Donc $\vec{a} \wedge \vec{b} = -\vec{b} \wedge \vec{a}$. ■

4° $(\lambda\vec{a}) \wedge \vec{b} = \lambda(\vec{a} \wedge \vec{b})$ (*associativité par rapport au facteur scalaire*).

Démonstration. La propriété est évidente si $\vec{a}$ et $\vec{b}$ sont colinéaires ou si $\lambda = 0$. Supposons que $\vec{a}$ et $\vec{b}$ ne sont pas colinéaires et que $\lambda \neq 0$. De la définition du produit vectoriel il s'ensuit que $|\lambda(\vec{a} \wedge \vec{b})| = |\lambda|\,|\vec{a}|\,|\vec{b}|\sin\varphi$ et $|(\lambda\vec{a}) \wedge \vec{b}| = |\lambda|\,|\vec{a}|\,|\vec{b}|\sin\varphi$, donc les vecteurs $(\lambda\vec{a}) \wedge \vec{b}$ et $\lambda(\vec{a} \wedge \vec{b})$ ont même module. Par ailleurs, ils sont parallèles, car orthogonaux aux vecteurs $\vec{a}$ et $\vec{b}$. Enfin ils sont de même sens (fig. 142) (pour $\lambda > 0$ ceci est évident, puisque les vecteurs $\lambda\vec{a}$ et $\vec{a}$ sont de

même sens ; pour $\lambda < 0$ les vecteurs $\lambda\vec{a}$ et $\vec{a}$ sont de sens contraires, donc le vecteur $(\lambda\vec{a}) \wedge \vec{b}$ est de sens contraire au vecteur $\vec{a} \wedge \vec{b}$, mais le vecteur $\lambda(\vec{a} \wedge \vec{b})$ est lui aussi de sens contraire au vecteur $\vec{a} \wedge \vec{b}$, donc les vecteurs $(\lambda\vec{a}) \wedge \vec{b}$ et $\lambda(\vec{a} \wedge \vec{b})$ sont de même sens). Par conséquent, les vecteurs $(\lambda\vec{a}) \wedge \vec{b}$ et $\lambda(\vec{a} \wedge \vec{b})$ sont égaux. ■

Utiliser les propriétés 3° et 4° pour prouver à titre d'exercice que $\vec{a} \wedge (\lambda\vec{b}) = \lambda(\vec{a} \wedge \vec{b})$.

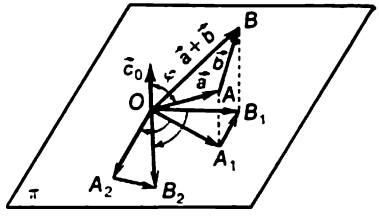


Fig. 143

5° $(\vec{a}+\vec{b}) \wedge \vec{c} = \vec{a} \wedge \vec{c} + \vec{b} \wedge \vec{c}$ (*distributivité par rapport à l'addition des vecteurs*).

Démonstration. Cette propriété est évidente si les vecteurs $\vec{a}$ et $\vec{b}$ sont colinéaires à $\vec{c}$ ou si l'un au moins des vecteurs $\vec{a}$, $\vec{b}$ et $\vec{c}$ est nul. Pour établir cette propriété dans les autres cas considérons un vecteur unitaire $\vec{c}_0$ de même direction et de même sens que $\vec{c}$ représenté par le vecteur $\overrightarrow{OC}$, où O est un point arbitraire de l'espace. Menons par O un plan π perpendiculaire à $\overrightarrow{OC}$ et construisons $\overrightarrow{OA} = \vec{a}$, $\overrightarrow{AB} = \vec{b}$ et $\overrightarrow{OB} = \vec{a} + \vec{b}$ (fig. 143). Soit OA_1B_1 le projeté du triangle OAB sur le plan π (si le point A_1 est sur la droite OB_1, le triangle OA_1B_1 dégénère en un segment). Soit OA_2B_2 le transformé du triangle OA_1B_1 par une rotation de $\pi/2$ autour de $\overrightarrow{OC}$ dans le sens rétrograde. Désignons par φ l'angle des vecteurs $\vec{c}_0$ et $\vec{a}$. Supposons pour fixer les idées que $0 < \varphi < \pi/2$ (comme sur la figure 143). Les autres cas de l'angle φ se traitent de façon analogue.

Considérons le vecteur $\overrightarrow{OA}_2$. Son module est $|\overrightarrow{OA}_2| = |\overrightarrow{OA}_1| = |\vec{a}| \cos(\pi/2 - \varphi) = |\vec{a}|\,|\vec{c}_0| \sin\varphi$, puisque $|\vec{c}_0| = 1$. Par ailleurs $\overrightarrow{OA}_2 \perp \vec{c}_0$, $\overrightarrow{OA}_2 \perp \vec{a}$ et les vecteurs $\vec{a}$, $\vec{c}_0$, $\overrightarrow{OA}_2$ forment un trièdre direct. Donc, par définition du produit vectoriel,

$$\overrightarrow{OA}_2 = \vec{a} \wedge \vec{c}_0.$$

En effectuant les mêmes raisonnements pour chaque vecteur $\overrightarrow{OB}_2$ et $\overrightarrow{A_2B_2}$, on obtient

$$\overrightarrow{OB}_2 = (\vec{a}+\vec{b}) \wedge \vec{c}_0, \quad \overrightarrow{A_2B_2} = \vec{b} \wedge \vec{c}_0.$$

Mais comme $\overrightarrow{OB_2} = \overrightarrow{OA_2} + \overrightarrow{A_2B_2}$, on a

$$(\vec{a}+\vec{b}) \wedge \vec{c}_0 = \vec{a} \wedge \vec{c}_0 + \vec{b} \wedge \vec{c}_0. \qquad (1)$$

Le vecteur $\vec{c}$ est de même sens que $\vec{c}_0$. Donc $\vec{c} = |\vec{c}|\,\vec{c}_0$. En multipliant les deux membres de l'égalité (1) par le nombre $|\vec{c}|$, on trouve $|\vec{c}|\,[(\vec{a}+\vec{b}) \wedge \vec{c}_0] = |\vec{c}|\,(\vec{a} \wedge \vec{c}_0 + \vec{b} \wedge \vec{c}_0)$. D'où, en vertu de la propriété 4°, $(\vec{a}+\vec{b}) \wedge |\vec{c}|\,\vec{c}_0 = \vec{a} \wedge |\vec{c}|\,\vec{c}_0 + \vec{b} \wedge |\vec{c}|\,\vec{c}_0$. En remplaçant $|\vec{c}|\,\vec{c}_0$ par $\vec{c}$, on trouve en définitive $(\vec{a}+\vec{b}) \wedge \vec{c} = \vec{a} \wedge \vec{c} + \vec{b} \wedge \vec{c}$. ■

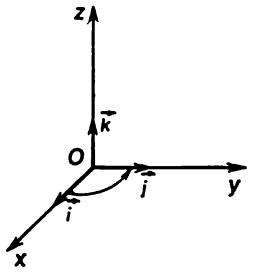


Fig. 144

Remarque 1. Cette propriété nous permet de multiplier vectoriellement les expressions vectorielles terme à terme et la propriété 4°, de regrouper les coefficients des facteurs vectoriels. Par exemple

$$(2\vec{a}+3\vec{b}) \wedge (4\vec{c}+5\vec{d}) = (2\vec{a}+3\vec{b}) \wedge 4\vec{c} + (2\vec{a}+3\vec{b}) \wedge 5\vec{d} =$$
$$= 2\vec{a} \wedge 4\vec{c} + 3\vec{b} \wedge 4\vec{c} + 2\vec{a} \wedge 5\vec{d} + 3\vec{b} \wedge 5\vec{d} =$$
$$= 8\,(\vec{a} \wedge \vec{c}) + 12\,(\vec{b} \wedge \vec{c}) + 10\,(\vec{a} \wedge \vec{d}) + 15\,(\vec{b} \wedge \vec{d}).$$

Signalons cependant qu'il faut rigoureusement respecter l'ordre des facteurs d'un produit vectoriel et en permutant ces facteurs il ne faut pas oublier de changer le signe. Exemple :

$$(\vec{a}+\vec{b}+\vec{c}) \wedge (2\vec{b}+3\vec{c}) = 2\,(\vec{a} \wedge \vec{b}) + 2\,(\vec{b} \wedge \vec{b}) + 2\,(\vec{c} \wedge \vec{b}) +$$
$$+ 3\,(\vec{a} \wedge \vec{c}) + 3\,(\vec{b} \wedge \vec{c}) + 3\,(\vec{c} \wedge \vec{c}) = 2\,(\vec{a} \wedge \vec{b}) - 2\,(\vec{b} \wedge \vec{c}) +$$
$$+ 3\,(\vec{a} \wedge \vec{c}) + 3\,(\vec{b} \wedge \vec{c}) = 2\,(\vec{a} \wedge \vec{b}) + (\vec{b} \wedge \vec{c}) + 3\,(\vec{a} \wedge \vec{c}).$$

Remarque 2. D'après la définition et les propriétés 1° et 3° du produit vectoriel, on a les relations suivantes pour les vecteurs de base $\vec{i}$, $\vec{j}$ et $\vec{k}$ (fig. 144)

$$\begin{array}{lll} \vec{i} \wedge \vec{i} = 0; & \vec{i} \wedge \vec{j} = \vec{k}; & \vec{i} \wedge \vec{k} = -\vec{j}; \\ \vec{j} \wedge \vec{i} = -\vec{k}; & \vec{j} \wedge \vec{j} = 0; & \vec{j} \wedge \vec{k} = \vec{i}; \\ \vec{k} \wedge \vec{i} = \vec{j}; & \vec{k} \wedge \vec{j} = -\vec{i}; & \vec{k} \wedge \vec{k} = 0. \end{array} \qquad (2)$$

3. Expression analytique du produit vectoriel.

Théorème 9.7. *Le produit vectoriel des vecteurs* $\vec{a} = \{X_1; Y_1; Z_1\}$ *et* $\vec{b} = \{X_2; Y_2; Z_2\}$ *s'exprime par la formule*

$$\vec{a} \wedge \vec{b} = \{(Y_1Z_2 - Y_2Z_1);\ (X_2Z_1 - X_1Z_2);\ (X_1Y_2 - X_2Y_1)\}.$$

En se servant des déterminants d'ordre deux *) on peut écrire cette formule sous la forme

$$\vec{a} \wedge \vec{b} = \left\{ \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix};\ \begin{vmatrix} Z_1 & X_1 \\ Z_2 & X_2 \end{vmatrix};\ \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix} \right\}.$$

Démonstration. Décomposons les vecteurs $\vec{a}$ et $\vec{b}$ suivant la base $\vec{i}, \vec{j}, \vec{k}$:

$$\vec{a} = X_1\vec{i} + Y_1\vec{j} + Z_1\vec{k}, \quad \vec{b} = X_2\vec{i} + Y_2\vec{j} + Z_2\vec{k}.$$

La remarque 1 nous donne

$$\vec{a} \wedge \vec{b} = X_1X_2(\vec{i} \wedge \vec{i}) + X_1Y_2(\vec{i} \wedge \vec{j}) + X_1Z_2(\vec{i} \wedge \vec{k}) + Y_1X_2(\vec{j} \wedge \vec{i}) +$$
$$+ Y_1Y_2(\vec{j} \wedge \vec{j}) + Y_1Z_2(\vec{j} \wedge \vec{k}) + Z_1X_2(\vec{k} \wedge \vec{i}) + Z_1Y_2(\vec{k} \wedge \vec{j}) + Z_1Z_2(\vec{k} \wedge \vec{k}).$$

D'où il vient en vertu de (2)

$$\vec{a} \wedge \vec{b} = (Y_1Z_2 - Y_2Z_1)\vec{i} + (X_2Z_1 - X_1Z_2)\vec{j} + (X_1Y_2 - X_2Y_1)\vec{k}$$

ou

$$\vec{a} \wedge \vec{b} = \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix}\vec{i} + \begin{vmatrix} Z_1 & X_1 \\ Z_2 & X_2 \end{vmatrix}\vec{j} + \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix}\vec{k}.$$

Ceci est la décomposition du vecteur $\vec{a} \wedge \vec{b}$ suivant la base $\vec{i}, \vec{j}, \vec{k}$; les coefficients de cette décomposition sont les composantes du vecteur $\vec{a} \wedge \vec{b}$. Donc

$$\vec{a} \wedge \vec{b} = \{X; Y; Z\}, \tag{3}$$

où

$$X = \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix}; \quad Y = \begin{vmatrix} Z_1 & X_1 \\ Z_2 & X_2 \end{vmatrix}; \quad Z = \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix}. \ \blacksquare$$

Exemple. Trouver les composantes du produit vectoriel des vecteurs $\vec{a} = \{2; 5; 7\}$ et $\vec{b} = \{1; 2; 4\}$.

---

*) On appelle *déterminant d'ordre deux* et on note $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$ le nombre égal à $a_1b_2 - a_2b_1$.

Solution. D'après la formule (3)

$$X=\begin{vmatrix} 5 & 7 \\ 2 & 4 \end{vmatrix}=6;\quad Y=\begin{vmatrix} 7 & 2 \\ 4 & 1 \end{vmatrix}=-1;\quad Z=\begin{vmatrix} 2 & 5 \\ 1 & 2 \end{vmatrix}=-1.$$

Donc $\vec{a}\wedge\vec{b}=\{6;\ -1;\ -1\}$.

§ 8. Produit mixte de trois vecteurs

1. Définition et interprétation géométrique du produit mixte.

Définition. On appelle *produit mixte* de trois vecteurs $\vec{a}$, $\vec{b}$ et $\vec{c}$ le nombre égal au produit scalaire du vecteur $\vec{a}$ par le produit vectoriel des vecteurs $\vec{b}$ et $\vec{c}$, c'est-à-dire

$$\vec{a}\cdot(\vec{b}\wedge\vec{c}).$$

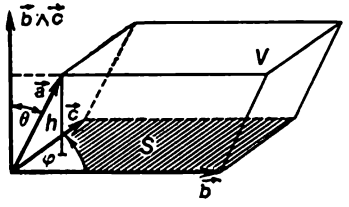


Fig. 145

La signification géométrique du produit mixte est donnée par e théorème suivant.

Théorème 9.8. *Le produit mixte* $\vec{a}\cdot(\vec{b}\wedge\vec{c})$ *est égal au volume v du parallélépipède construit sur les vecteurs* $\vec{a}$, $\vec{b}$, $\vec{c}$ *pris avec le signe « + » si le trièdre* $\vec{a}$, $\vec{b}$, $\vec{c}$ *est direct et avec le signe « — » s'il est rétrograde. Si* $\vec{a}$, $\vec{b}$ *et* $\vec{c}$ *sont coplanaires,* $\vec{a}\cdot(\vec{b}\wedge\vec{c})=0$. *Autrement dit,*

$$\vec{a}\cdot(\vec{b}\wedge\vec{c})=\begin{cases} v \text{ si } \vec{a},\ \vec{b},\ \vec{c} \text{ est un trièdre direct,} \\ -v \text{ si } \vec{a},\ \vec{b},\ \vec{c} \text{ est rétrograde,} \\ 0 \text{ si } \vec{a},\ \vec{b} \text{ et } \vec{c} \text{ sont coplanaires.} \end{cases}$$

Démonstration. Soient $\vec{a}$, $\vec{b}$ et $\vec{c}$ des vecteurs formant un trièdre direct. Désignons par v le volume du parallélépipède construit sur ces vecteurs, par s l'aire du parallélogramme construit sur les vecteurs $\vec{b}$ et $\vec{c}$ et par h la hauteur du parallélépipède (fig. 145). Par définition du produit scalaire et du produit vectoriel, on a

$$\vec{a}\cdot(\vec{b}\wedge\vec{c})=|\vec{a}|\ |\vec{b}\wedge\vec{c}|\cos\theta=|\vec{a}|\ |\vec{b}|\ |\vec{c}|\sin\varphi\cos\theta=$$
$$=|\vec{b}|\ |\vec{c}|\sin\varphi\ |\vec{a}|\cos\theta,$$

où φ est l'angle des vecteurs $\vec{b}$ et $\vec{c}$, θ l'angle des vecteurs $\vec{a}$ et $\vec{b}\wedge\vec{c}$. Puisque $|\vec{b}|\,|\vec{c}|\sin\varphi=s$ et $|\vec{a}|\cos\theta=h$, il vient

$\vec{a}\cdot(\vec{b}\wedge\vec{c}) = sh = v$. Si le trièdre $\vec{a}$, $\vec{b}$, $\vec{c}$ est rétrograde, on a $h = |\vec{a}|\cos(\pi-\theta) = -|\vec{a}|\cos\theta$. Donc $\vec{a}\cdot(\vec{b}\wedge\vec{c}) = -sh = -v$. Ce qui prouve la première proposition.

Prouvons la deuxième proposition. Supposons que les vecteurs $\vec{a}$, $\vec{b}$ et $\vec{c}$ sont coplanaires. Si $\vec{a} = \vec{0}$, il est évident que $\vec{a}\cdot(\vec{b}\wedge\vec{c}) = 0$. Supposons que $\vec{a} \neq \vec{0}$. Alors ou bien $\vec{b}\wedge\vec{c} = \vec{0}$ (si $\vec{b}$ et $\vec{c}$ sont colinéaires), ou bien $(\vec{b}\wedge\vec{c}) \perp \vec{a}$ (si $\vec{b}$ et $\vec{c}$ ne sont pas colinéaires). Dans tous les cas $\vec{a}\cdot(\vec{b}\wedge\vec{c}) = 0$. ■

On a ainsi prouvé que si les vecteurs $\vec{a}$, $\vec{b}$ et $\vec{c}$ sont coplanaires, $\vec{a}\cdot(\vec{b}\wedge\vec{c}) = 0$. La réciproque est vraie: si $\vec{a}\cdot(\vec{b}\wedge\vec{c}) = 0$, les vecteurs $\vec{a}$, $\vec{b}$ et $\vec{c}$ sont coplanaires. En effet, si les vecteurs $\vec{a}$, $\vec{b}$ et $\vec{c}$ n'étaient pas coplanaires, le produit mixte $\vec{a}\cdot(\vec{b}\wedge\vec{c}) = \pm v \neq 0$ d'après le théorème 9.8, ce qui est contraire à l'hypothèse.

Corollaire. *Le théorème nous permet d'établir immédiatement l'identité*

$$\vec{a}\cdot(\vec{b}\wedge\vec{c}) = (\vec{a}\wedge\vec{b})\cdot\vec{c}, \tag{1}$$

qui exprime que les signes $\cdot$ et $\wedge$ peuvent être échangés dans le produit mixte.

En effet, d'après la propriété 1° du produit scalaire,

$$(\vec{a}\wedge\vec{b})\cdot\vec{c} = \vec{c}\cdot(\vec{a}\wedge\vec{b}). \tag{2}$$

D'après le théorème 9.8,

$$\vec{a}\cdot(\vec{b}\wedge\vec{c}) = \pm v, \quad \vec{c}\cdot(\vec{a}\wedge\vec{b}) = \pm v. \tag{3}$$

Les trièdres $\vec{a}$, $\vec{b}$, $\vec{c}$ et $\vec{c}$, $\vec{a}$, $\vec{b}$ étant de même sens, c'est-à-dire ou tous deux directs ou tous deux rétrogrades, les seconds membres des égalités (3) doivent être pris avec le même signe d'après le théorème 9.8. On a donc

$$\vec{a}\cdot(\vec{b}\wedge\vec{c}) = \vec{c}\cdot(\vec{a}\wedge\vec{b})$$

et en vertu de l'égalité (2)

$$\vec{a}\cdot(\vec{b}\wedge\vec{c}) = (\vec{a}\wedge\vec{b})\cdot\vec{c},$$

ce qu'on voulait.

D'après l'identité (1) les produits mixtes $\vec{a}\cdot(\vec{b}\wedge\vec{c})$ et $(\vec{a}\wedge\vec{b})\cdot\vec{c}$ peuvent être désignés plus simplement par le symbole $\vec{a}\vec{b}\vec{c}$.

2. Expression analytique du produit mixte.

Théorème 9.9. *Le produit mixte* $\vec{a}\vec{b}\vec{c}$ *des vecteurs* $\vec{a} = \{X_1;\ Y_1;\ Z_1\}$, $\vec{b} = \{X_2;\ Y_2;\ Z_2\}$, $\vec{c} = \{X_3;\ Y_3;\ Z_3\}$ *s'exprime par la formule*

$$\vec{a}\vec{b}\vec{c} = X_1\begin{vmatrix} Y_2 & Z_2 \\ Y_3 & Z_3 \end{vmatrix} + Y_1\begin{vmatrix} Z_2 & X_2 \\ Z_3 & X_3 \end{vmatrix} + Z_1\begin{vmatrix} X_2 & Y_2 \\ X_3 & Y_3 \end{vmatrix}.$$

Démonstration. On a $\vec{a}\vec{b}\vec{c} = \vec{a}\cdot(\vec{b}\wedge\vec{c})$. D'après le théorème 9.7,

$$\vec{b}\wedge\vec{c} = \left\{\begin{vmatrix} Y_2 & Z_2 \\ Y_3 & Z_3 \end{vmatrix};\ \begin{vmatrix} Z_2 & X_2 \\ Z_3 & X_3 \end{vmatrix};\ \begin{vmatrix} X_2 & Y_2 \\ X_3 & Y_3 \end{vmatrix}\right\}.$$

En multipliant scalairement le vecteur $\vec{a}$ par le vecteur $\vec{b}\wedge\vec{c}$ et en appliquant le théorème 9.6, on obtient

$$\vec{a}\vec{b}\vec{c} = X_1\begin{vmatrix} Y_2 & Z_2 \\ Y_3 & Z_3 \end{vmatrix} + Y_1\begin{vmatrix} Z_2 & X_2 \\ Z_3 & X_3 \end{vmatrix} + Z_1\begin{vmatrix} X_2 & Y_2 \\ X_3 & Y_3 \end{vmatrix}. \blacksquare$$

Exemple. Trouver le volume du tétraèdre $ABCD$ de sommets A (1; 1; 1), B (4; 4; 4), C (3; 5; 5) et D (2; 4; 7).

Solution. On sait du cours de géométrie élémentaire que le volume v_t du tétraèdre $ABCD$ est égal au sixième du volume du parallélépipède construit sur les vecteurs $\vec{AB}$, $\vec{AC}$ et $\vec{AD}$; on en déduit grâce au théorème 9.8 que v_t est égal au 1/6 de la valeur absolue du produit mixte $\vec{AB}\cdot\vec{AC}\cdot\vec{AD}$. Pour calculer ce produit mixte déterminons les composantes des vecteurs $\vec{AB}$, $\vec{AC}$ et $\vec{AD}$. D'après le théorème 9.2: $\vec{AB} = \{3;\ 3;\ 3\}$, $\vec{AC} = \{2;\ 4;\ 4\}$, $\vec{AD} = \{1;\ 3;\ 6\}$. Le théorème 9.8 nous donne

$$\vec{AB}\cdot\vec{AC}\cdot\vec{AD} = 3\begin{vmatrix} 4 & 4 \\ 3 & 6 \end{vmatrix} + 3\begin{vmatrix} 4 & 2 \\ 6 & 1 \end{vmatrix} + 3\begin{vmatrix} 2 & 4 \\ 1 & 3 \end{vmatrix} =$$

$$= 3\cdot 12 - 3\cdot 8 + 3\cdot 2 = 18.$$

D'où

$$v_t = \frac{1}{6}\cdot 18 = 3.$$

§ 9. Equations d'une surface et d'une ligne

Soient donnés un système de coordonnées rectangulaires $Oxyz$, une surface S (fig. 146) et une équation

$$F(x,\ y,\ z) = 0. \qquad (1)$$

On dira que l'équation (1) est l'*équation de la surface* S dans le système de coordonnées $Oxyz$ si elle est satisfaite par les coordon-

nées de tous les points $M(x; y; z)$ de S et de ces points seulement.

Du point de vue de cette définition, la surface S est l'ensemble des points dont les coordonnées vérifient l'équation (1).

Exemple. Dans un système de coordonnées rectangulaires, l'équation

$$x^2 + y^2 + z^2 - R^2 = 0 \quad \text{ou} \quad x^2 + y^2 + z^2 = R^2$$

définit une sphère de rayon R centrée en O (0, 0, 0) (fig. 147).

En effet, si M est un point arbitraire, la formule (7) (cf. § 2, n° 5) nous donne

$$OM = \sqrt{x^2+y^2+z^2}.$$

Donc cette équation est satisfaite par les coordonnées des points situés à une distance R de O et de ces points seulement. Par suite,

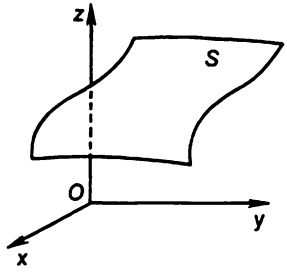


Fig. 146

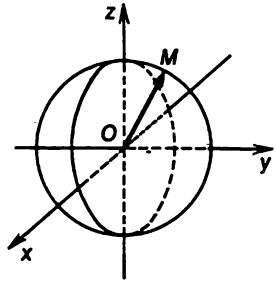


Fig. 147

l'ensemble des points dont les coordonnées vérifient cette équation est la sphère de rayon R et de centre O.

Dans l'espace, une ligne peut être envisagée comme l'intersection de deux surfaces, c'est-à-dire comme l'ensemble des points appartenant simultanément à deux surfaces, et définie respectivement par la donnée de deux équations. Ainsi les équations

$$\begin{cases} F_1(x; y; z) = 0, \\ F_2(x; y; z) = 0 \end{cases} \tag{2}$$

s'appellent *équations de la ligne* L si elles sont satisfaites par les coordonnées des points de L et de ces points seulement.

Exemple: les équations des sphères

$$\begin{cases} x^2+y^2+(z-3)^2=10, \\ x^2+y^2+z^2=1 \end{cases}$$

définissent le cercle du plan Oxy de rayon 1 centré en O.

§ 10. Equation d'un cylindre

Soient données une ligne L dans le plan Oxy et une direction fixe Δ dans l'espace (fig. 148). Une surface cylindrique S est la surface engendrée par une droite variable parallèle à la direction fixe Δ et s'appuyant sur la ligne L. La droite variable s'appelle *génératrice*, la ligne L, *directrice* de S.

Pour fixer les idées, considérons une surface cylindrique S de génératrice parallèle à Oz et prouvons qu'elle est définie par une équation de la forme

$$F(x;\ y) = 0. \tag{1}$$

En effet, supposons que l'équation (1) est l'équation de la directrice L. Prenons un point quelconque $M(x;\ y;\ z)$ de S. Ce

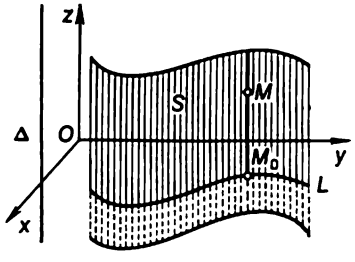


Fig. 148

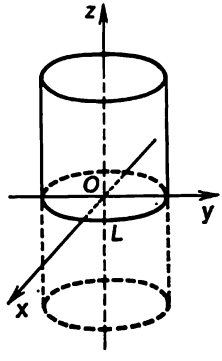


Fig. 149

point est situé sur une génératrice. Si M_0 est la trace de cette génératrice sur le plan Oxy, le point $M_0 \in L$ et ses coordonnées x et y vérifient l'équation (1). Les points x, y, z satisferont aussi à cette équation, puisque $F(x;\ y)$ ne dépend pas de z. Ainsi les coordonnées x, y, z de tout point $M \in S$ vérifient l'équation (1). Il est évident que si $M(x;\ y;\ z) \notin S$, alors $M_0(x;\ y) \notin L$, c'est-à-dire que les coordonnées x, y ne vérifient pas l'équation (1). Ceci prouve que (1) est l'équation de la surface S.

Donc l'équation d'un cylindre engendré par une génératrice parallèle à l'axe Oz ne contient pas la coordonnée z et coïncide avec l'équation de la directrice. Par exemple, si la directrice est l'ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{2}$$

le cylindre engendré s'appelle *cylindre elliptique* et son équation est l'équation (2).

Remarquons que l'équation $F(x; y) = 0$ définit la ligne L dans le plan Oxy mais que dans l'espace cette ligne est donnée par deux équations

$$\begin{cases} F(x; y) = 0, \\ z = 0. \end{cases}$$

Ainsi par exemple, l'équation $x^2 + y^2 - R^2 = 0$ définit un cylindre circulaire (fig. 149) dont la directrice, qui est un cercle du plan Oxy, est donnée par les équations

$$\begin{cases} x^2 + y^2 = R^2 \\ z = 0. \end{cases}$$

§ 11. Equations du plan

Montrons que les plans et eux seuls sont des surfaces du premier ordre et considérons deux types d'équations du plan.

1. Equation générale du plan. Soient donnés un système de coordonnées rectangulaires $Oxyz$ un plan arbitraire π, un point $M_0(x_0; y_0; z_0) \in \pi$, un vecteur $\vec{N} = \{A; B; C\}$ orthogonal à π (fig. 150).

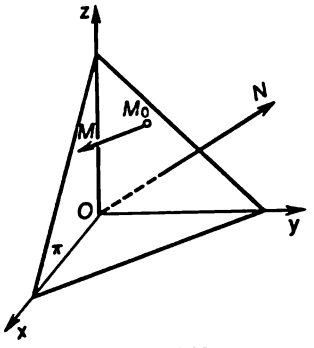


Fig. 150

Considérons un point arbitraire $M(x; y; z)$. Le point M appartient au plan π si et seulement si les vecteurs $\overrightarrow{M_0M}$ et $\vec{N}$ sont orthogonaux. Les composantes du vecteur $\overrightarrow{M_0M}$ étant égales à $x - x_0$, $y - y_0$, $z - z_0$, la condition d'orthogonalité de deux vecteurs [cf. § 6, formule (3)] nous dit que le point $M(x; y; z)$ appartient au plan π si et seulement si

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \qquad (1)$$

Ceci est l'équation cherchée du plan π, puisqu'elle est vérifiée par les coordonnées $x; y; z$ des points M de π, et de ces points seulement.

En chassant les parenthèses on ramène l'équation (1) à la forme

$$Ax + By + Cz + (-Ax_0 - By_0 - Cz_0) = 0.$$

En désignant le nombre $-Ax_0 - By_0 - Cz_0$ par D, on obtient

$$Ax + By + Cz + D = 0. \qquad (2)$$

L'équation (2) s'appelle *équation générale du plan*. Donc le plan est une surface du premier ordre, c'est-à-dire une surface dont l'équation est un polynôme du premier degré en x, y, z.

Réciproque : toute équation du premier degré de la forme (2) définit un plan dans un système de coordonnées rectangulaires.

En effet, soient donnés un système de coordonnées rectangulaires $Oxyz$ et une équation $Ax + By + Cz + D = 0$ de coefficients A, B, C, D, arbitraires, A, B et C étant non tous nuls. Cette équation admet visiblement au moins une racine x_0, y_0, z_0 $\left(\text{si, par exemple, } C \neq 0, \text{ en choisissant arbitrairement } x_0 \text{ et } y_0, \text{ on obtient } z_0 = -\frac{A}{C}x_0 - \frac{B}{C}y_0 - \frac{D}{C}\right)$. Donc il existe au moins un point $M_0(x_0; y_0; z_0)$ dont les coordonnées vérifient l'équation, c'est-à-dire que $Ax_0 + By_0 + Cz_0 + D = 0$. En retranchant cette égalité de l'équation $Ax + By + Cz + D = 0$, on obtient l'équation équivalente

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

Cette équation (donc l'équation $Ax + By + Cz + D = 0$) est confondue avec l'équation (1) ; elle définit par suite un plan π passant par $M_0(x_0; y_0; z_0)$ et orthogonal au vecteur $\vec{N} = \{A; B; C\}$.

Le vecteur $\vec{N} = \{A; B; C\}$ orthogonal au plan π s'appelle *vecteur normal* à π.

Exemple. Etablir l'équation du plan passant par le point $M_0(1; 1; 1)$ et orthogonal au vecteur $\vec{N} = \{2; 2; 3\}$.

Solution. D'après la formule (1) l'équation cherchée est

$$2(x - 1) + 2(y - 1) + 3(z - 1) = 0$$

ou

$$2x + 2y + 3z - 7 = 0.$$

Prouvons en conclusion le théorème suivant.

Théorème 9.10. *Si deux équations* $A_1x + B_1y + C_1z + D_1 = 0$ *et* $A_2x + B_2y + C_2z + D_2 = 0$ *définissent un même plan, leurs coefficients sont proportionnels.*

Démonstration. En effet, les vecteurs $\vec{N}_1 = \{A_1; B_1; C_1\}$ et $\vec{N}_2 = \{A_2; B_2; C_2\}$ sont orthogonaux à ce plan, donc parallèles. Donc les nombres A_1, B_1 et C_1 sont respectivement proportionnels aux nombres A_2, B_2 et C_2 (cf. formule (2), § 4), c'est-à-dire que

$$\frac{A_2}{A_1} = \frac{B_2}{B_1} = \frac{C_2}{C_1} = \mu$$

ou $A_2 = \mu A_1$, $B_2 = \mu B_1$, $C_2 = \mu C_1$. En multipliant la première équation par μ et en retranchant de la seconde, on obtient

$D_2 - \mu D_1 = 0$, c'est-à-dire que $D_2 = \mu D_1$ et par suite

$$\frac{A_2}{A_1} = \frac{B_2}{B_1} = \frac{C_2}{C_1} = \frac{D_2}{D_1} = \mu. \blacksquare$$

2. Angle de deux plans. Considérons deux plans π_1 et π_2 d'équations respectives $A_1x + B_1y + C_1z + D_1 = 0$ et $A_2x + B_2y + C_2z + D_2 = 0$.

L'angle φ des plans π_1 et π_2 est l'ensemble $\{\varphi, \varphi + \pi\}$, où φ est égal à l'angle de leurs vecteurs normaux $\vec{N}_1 = \{A_1; B_1; C_1\}$ et $\vec{N}_2 = \{A_2; B_2; C_2\}$. Cet angle est donné par la formule

$$\cos\varphi = \frac{\vec{N}_1 \cdot \vec{N}_2}{|N_1|\,|\vec{N}_2|} = \frac{A_1A_2 + B_1B_2 + C_1C_2}{\sqrt{A_1^2 + B_1^2 + C_1^2}\,\sqrt{A_2^2 + B_2^2 + C_2^2}}. \qquad (3)$$

3. Condition de parallélisme des plans. Si des plans π_1 et π_2 sont parallèles, leurs vecteurs normaux $\vec{N}_1$ et $\vec{N}_2$ le sont aussi, et réciproquement. Donc

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}. \qquad (4)$$

La condition (4) est une condition nécessaire et suffisante pour que les plans π_1 et π_2 soient parallèles.

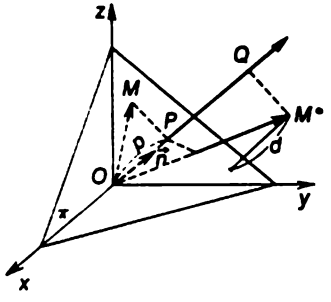


Fig. 151

4. Condition de perpendicularité des plans. Si des plans π_1 et π_2 sont perpendiculaires, leurs vecteurs normaux $\vec{N}_1$ et $\vec{N}_2$ sont orthogonaux, et réciproquement. La formule (3) nous donne immédiatement la condition nécessaire et suffisante de perpendicularité des plans π_1 et π_2:

$$A_1A_2 + B_1B_2 + C_1C_2 = 0.$$

5. Equation normale du plan. Distance d'un point à un plan. Soit donné un plan π dans l'espace rapporté à un système de coordonnées rectangulaires $Oxyz$ (fig. 151). Menons par l'origine O une droite perpendiculaire au plan π. Cette droite s'appelle *normale*. Désignons par P la trace de la normale sur le plan π. Orientons cette normale de O vers P. Si les points O et P sont confondus, le sens sera arbitrairement choisi. Soient α, β et γ les angles de la normale orientée et des axes de coordonnées, p la mesure algébrique de $\overrightarrow{OP}$.

Etablissons l'équation du plan π en admettant que $\cos\alpha$, $\cos\beta$, $\cos\gamma$ et p sont connus. Considérons à cet effet le vecteur unitaire $\vec{n}$ de la normale.

On a

$$\vec{n} = \{\cos\alpha;\ \cos\beta;\ \cos\gamma\}. \tag{5}$$

Soit $M(x; y; z)$ un point arbitraire. Ce point appartient au plan π si et seulement si la mesure algébrique du projeté de $\overrightarrow{OM}$ sur la normale est égale à p, c'est-à-dire si

$$\overline{\mathrm{pr}_{\vec{n}}\overrightarrow{OM}} = p. \tag{6}$$

Remarquons maintenant que $\overline{\mathrm{pr}_{\vec{n}}\overrightarrow{OM}} = \vec{n}\cdot\overrightarrow{OM}$ et $\overrightarrow{OM} = \{x; y; z\}$. Le théorème 9.6 nous donne, compte tenu de (5),

$$\overline{\mathrm{pr}_{\vec{n}}\overrightarrow{OM}} = \vec{n}\cdot\overrightarrow{OM} = x\cos\alpha + y\cos\beta + z\cos\gamma. \tag{7}$$

Des égalités (6) et (7) il s'ensuit que le point $M(x; y; z)$ est sur le plan π si et seulement si ses coordonnées satisfont à l'équation

$$x\cos\alpha + y\cos\beta + z\cos\gamma - p = 0, \tag{8}$$

qui est l'équation cherchée. Cette équation s'appelle *équation normale du plan.*

Théorème 9.11. *La distance d d'un point $M^*(x^*; y^*; z^*)$ à un plan d'équation normale*

$$x\cos\alpha + y\cos\beta + z\cos\gamma - p = 0$$

est définie par la formule

$$d = |\, x^*\cos\alpha + y^*\cos\beta + z^*\cos\gamma - p\,|.$$

Démonstration. Soit Q le projeté de M^* sur la normale orientée (fig. 151). D'après la relation de Chasles, $\overline{PQ} = \overline{OQ} - \overline{OP}$, d'où

$$d = |\,\overline{PQ}\,| = |\,\overline{OQ} - \overline{OP}\,|.$$

Or $\overline{OQ} = \overline{\mathrm{pr}_{\vec{n}}\overrightarrow{OM}^*}$, $\overline{OP} = p$, donc

$$d = \left|\overline{\mathrm{pr}_{\vec{n}}\overrightarrow{OM}^*} - p\right|. \tag{9}$$

D'autre part, $\overline{\mathrm{pr}_{\vec{n}}\overrightarrow{OM^*}} = \vec{n}\cdot\overrightarrow{OM^*}$. Le théorème 9.5 nous donne, compte tenu de (5),

$$\overline{\mathrm{pr}_{\vec{n}}\overrightarrow{OM^*}} = \vec{n}\cdot\overrightarrow{OM^*} = x^*\cos\alpha + y^*\cos\beta + z^*\cos\gamma. \qquad (10)$$

En combinant (9) et (10) on obtient en définitive

$$d = |\, x^*\cos\alpha + y^*\cos\beta + z^*\cos\gamma - p\,|. \blacksquare$$

Montrons maintenant comment ramener l'équation générale du plan à la forme normale. Soit

$$Ax + By + Cz + D = 0 \qquad (11)$$

l'équation générale d'un plan et

$$x\cos\alpha + y\cos\beta + z\cos\gamma - p = 0. \qquad (12)$$

son équation normale. Les équations (11) et (12) définissant le même plan, leurs coefficients sont proportionnels en vertu du théorème 9.10. Ceci signifie qu'en multipliant (11) par un facteur $\mu \neq 0$, on obtient l'équation

$$\mu Ax + \mu By + \mu Cz + \mu D = 0$$

qui est confondue avec (12), d'où

$$\mu A = \cos\alpha, \quad \mu B = \cos\beta, \quad \mu C = \cos\gamma, \quad \mu D = -p. \qquad (13)$$

Pour trouver le facteur μ, élevons les trois premières égalités (13) au carré et additionnons; on obtient alors

$$\mu^2 (A^2 + B^2 + C^2) = \cos^2\alpha + \cos^2\beta + \cos^2\gamma.$$

Le second membre de cette égalité est égal à 1 d'après la formule (6) du § 2. Donc

$$\mu = \pm \frac{1}{\sqrt{A^2+B^2+C^2}}.$$

Le nombre μ qui sert à ramener l'équation générale du plan à la forme normale s'appelle *facteur de normalisation* de cette équation. Le signe de μ est défini par l'égalité $\mu D = -p$, c'est-à-dire que μ est de signe contraire à D.

Si $D = 0$ dans l'équation (11), μ est pris indifféremment avec le signe + ou le signe —.

Exemple. Trouver la distance d du point M^* (1 ; 1 ; 1) au plan $x + 2y + 2z - 8 = 0$.

S o l u t i o n. Pour appliquer le théorème 9.11, il faut d'abord ramener cette équation à la forme normale. Le facteur de normalisation est

$$\mu = 1/\sqrt{1^2 + 2^2 + 2^2} = 1/3.$$

En multipliant l'équation donnée par 1/3 on obtient l'équation normale du plan

$$\frac{1}{3}x + \frac{2}{3}y + \frac{2}{3}z - \frac{8}{3} = 0.$$

En portant les coordonnées de M^* dans cette équation on trouve

$$d = |\,1/3 + 2/3 + 2/3 - 8/3\,| = |\,-3/3\,| = 1.$$

§ 12. Equations de la droite

Comme déjà signalé, une ligne peut être envisagée comme l'intersection de deux surfaces dans l'espace et définie par la donnée de deux équations. En particulier, chaque droite peut être considérée comme l'intersection de deux plans et représentée par conséquent par deux équations du premier degré.

Supposons que l'espace est rapporté à un système de coordonnées rectangulaires $Oxyz$ et soit L une droite arbitraire. Désignons par π_1 et π_2 deux plans distincts d'équations respectives

$$A_1x + B_1y + C_1z + D_1 = 0, \quad A_2x + B_2y + C_2z + D_2 = 0 \quad (1)$$

se coupant suivant la droite L. Les équations (1) représentent la droite L si et seulement si les plans π_1 et π_2 ne sont ni parallèles, ni confondus, c'est-à-dire si les vecteurs normaux $\vec{N}_1 = \{A_1; B_1; C_1\}$ et $\vec{N}_2 = \{A_2; B_2; C_2\}$ de ces plans ne sont pas colinéaires (les coefficients A_1, B_1, C_1 ne sont pas proportionnels aux coefficients A_2, B_2, C_2).

Les équations (1) s'appellent *équations générales de la droite.*

1. Equations canoniques de la droite. Les équations (1) ne sont pas toujours commodes à l'usage, c'est pourquoi elles sont utilisées sous une autre forme que nous proposons d'établir.

Soient donnés une droite L et un vecteur $\vec{a}$ non nul porté par L ou parallèle à elle. Le vecteur $\vec{a}$ s'appelle *vecteur directeur de L.* Etablissons l'équation de la droite de vecteur directeur $\vec{a} = \{l; m; n\}$ passant par un point donné $M_0(x_0; y_0; z_0)$ (fig. 152).

Soit $M(x; y; z)$ un point arbitraire. Ce point appartient à la droite cherchée si et seulement si le vecteur $\overrightarrow{M_0M} = \{x - x_0;$

$y - y_0; z - z_0\}$ est colinéaire au vecteur directeur $\vec{a} = \{l; m; n\}$, c'est-à-dire si et seulement si les composantes du vecteur $\overrightarrow{M_0M}$ sont proportionnelles à celles du vecteur $\vec{a}$:

$$\frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n}. \qquad (2)$$

Les équations (2) sont les équations cherchées. On les appelle *équations canoniques de la droite.*

Pour former les équations canoniques (2) si la droite L est représentée par les équations (1), il faut:

1) déterminer les coordonnées d'un point quelconque $M_0(x_0; y_0; z_0)$ de L; à cet effet il faut se donner une valeur

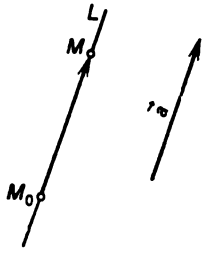


Fig. 152

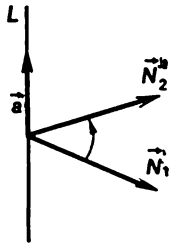


Fig. 153

arbitraire d'une coordonnée, substituer cette valeur dans les équations (1) et trouver les valeurs des deux autres coordonnées en résolvant le système d'équations (1);

2) trouver un vecteur directeur $\vec{a} = \{l; m; n\}$. La droite L étant l'intersection des plans π_1 et π_2, elle est orthogonale à leurs vecteurs normaux respectifs $\vec{N}_1$ et $\vec{N}_2$ (fig. 153). Donc pour vecteur $\vec{a}$ on peut prendre tout vecteur orthogonal à $\vec{N}_1$ et $\vec{N}_2$, par exemple leur produit vectoriel $\vec{a} = \vec{N}_1 \wedge \vec{N}_2$. Les composantes des vecteurs $\vec{N}_1$ et $\vec{N}_2$ étant connues: $\vec{N}_1 = \{A_1; B_1; C_1\}$ et $\vec{N}_2 = \{A_2; B_2; C_2\}$, le théorème 9.7 nous donne celles du vecteur $\vec{a}$:

$$\vec{a} = \left\{ \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}; \begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix}; \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \right\} = \{l; m; n\}.$$

Exemple 1. Trouver les équations canoniques de la droite

$$\begin{cases} 3x + 2y + 4z - 11 = 0, \\ 2x + y - 3z - 1 = 0. \end{cases}$$

S o l u t i o n. En prenant, par exemple, $x_0 = 1$, on obtient le système

$$\begin{cases} 2y_0 + 4z_0 - 8 = 0, \\ y_0 - 3z_0 + 1 = 0 \end{cases}$$

qui donne $y_0 = 2$, $z_0 = 1$. $M_0\,(1\,;\,2\,;\,1)$ est donc un point de la droite. Cherchons maintenant un vecteur directeur $\vec{a}$. On a $\vec{N}_1 = \{3\,;\,2\,;\,4\}$, $\vec{N}_2 = \{2\,;\,1\,;\,-3\}$, d'où $\vec{a} = \vec{N}_1 \wedge \vec{N}_2 = \{-10\,;\,17\,;\,-1\}$, c'est-à-dire que $l = -10$, $m = 17$, $n = -1$. Les équations canoniques cherchées sont donc:

$$\frac{x-1}{-10} = \frac{y-2}{17} = \frac{z-1}{-1}.$$

2. Equations paramétriques de la droite. On a parfois intérêt à se donner une droite sous une forme autre que la forme canonique. Soit L une droite représentée par des équations (2). Désignons par t la valeur commune de ces rapports. Alors

$$\frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n} = t,$$

d'où

$$x = x_0 + lt, \quad y = y_0 + mt, \quad z = z_0 + nt. \tag{3}$$

Les équations (3) s'appellent *équations paramétriques* de la droite L, de vecteur directeur $\vec{a} = \{l\,;\,m\,;\,n\}$, passant par le point $M_0\,(x_0\,;\,y_0\,;\,z_0)$. Dans les équations (3), t est traité comme un paramètre variant arbitrairement sur $]-\infty, +\infty[$ et x, y, z comme des fonctions de t. Lorsque t varie, le point $M\,(x\,;\,y\,;\,z)$ se déplace sur la droite donnée.

Les équations paramétriques sont commodes dans les cas où il faut déterminer la trace d'une droite sur un plan. En effet, soient donnés un plan π et une droite L non parallèle à π, d'équations respectives

$$Ax + By + Cz + D = 0 \quad \text{et}$$
$$x = x_0 + lt, \quad y = y_0 + mt, \quad z = z_0 + nt.$$

Pour déterminer la trace de la droite L sur le plan π, portons les expressions de x, y et z dans l'équation de π. Toutes transformations faites, on obtient

$$t = -\frac{Ax_0 + By_0 + Cz_0 + D}{Al + Bm + Cn},$$

le dénominateur étant non nul, puisque L n'est pas parallèle à π (cf. § 13). En portant la valeur de t trouvée dans l'équation de la droite, on détermine la trace $M\,(x\,;\,y\,;\,z)$ de la droite L sur le plan π.

3. Angles de deux droites. Soient L_1 et L_2 deux droites d'équations respectives

$$\frac{x-x_1}{l_1}=\frac{y-y_1}{m_1}=\frac{z-z_1}{n_1} \quad \text{et} \quad \frac{x-x_2}{l_2}=\frac{y-y_2}{m_2}=\frac{z-z_2}{n_2}.$$

L'angle de deux droites L_1 et L_2 de l'espace est l'ensemble $\{\varphi, \varphi+\pi\}$, où φ est l'angle de leurs vecteurs directeurs $\vec{a}_1=\{l_1; m_1; n_1\}$ et $\vec{a}_2=\{l_2; m_2; n_2\}$. Cet angle est donné par la formule :

$$\cos\varphi=\frac{l_1l_2+m_1m_2+n_1n_2}{\sqrt{l_1^2+m_1^2+n_1^2}\,\sqrt{l_2^2+m_2^2+n_2^2}}.$$

4. Condition de parallélisme de deux droites. Deux droites L_1 et L_2 sont parallèles si et seulement si leurs vecteurs directeurs $\vec{a}_1=\{l_1; m_1; n_1\}$ et $\vec{a}_2=\{l_2; m_2; n_2\}$ sont colinéaires, c'est-à-dire si et seulement si

$$\frac{l_2}{l_1}=\frac{m_2}{m_1}=\frac{n_2}{n_1}.$$

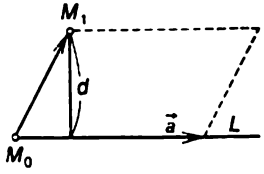


Fig. 154

5. Condition d'orthogonalité de deux droites. Deux droites L_1 et L_2 sont orthogonales si et seulement si leurs vecteurs directeurs $a_1=\{l_1; m_1; n_1\}$ et $\vec{a}_2=\{l_2; m_2; n_2\}$ sont orthogonaux, c'est-à-dire si et seulement si

$$l_1l_2+m_1m_2+n_1n_2=0.$$

6. Distance d'un point à une droite. Soient donnés dans l'espace un point $M_1(x_1; y_1; z_1)$ et une droite L d'équation $\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}$. La distance d de M_1 à L est la hauteur du parallélogramme construit sur les vecteurs $\overrightarrow{M_0M_1}=\{x_1-x_0; y_1-y_0; z_1-z_0\}$ et $\vec{a}=\{l; m; n\}$ (fig. 154).

Soit $\vec{p}=\overrightarrow{M_0M_1}\wedge\vec{a}$. Puisque $|\vec{p}|$ est égal à l'aire du parallélogramme construit sur les vecteurs $\overrightarrow{M_0M_1}$ et $\vec{a}$, il vient

$$d=\frac{|\vec{p}|}{|\vec{a}|}$$

où

$$|\vec{p}|=\sqrt{\begin{vmatrix} y_1-y_0 & z_1-z_0 \\ m & n \end{vmatrix}^2+\begin{vmatrix} z_1-z_0 & x_1-x_0 \\ n & l \end{vmatrix}^2+\begin{vmatrix} x_1-x_0 & y_1-y_0 \\ l & m \end{vmatrix}^2},$$

$$|\vec{a}|=\sqrt{l^2+m^2+n^2}.$$

Donc

$$d = \frac{\sqrt{\begin{vmatrix} y_1-y_0 & z_1-z_0 \\ m & n \end{vmatrix}^2 + \begin{vmatrix} z_1-z_0 & x_1-x_0 \\ n & l \end{vmatrix}^2 + \begin{vmatrix} x_1-x_0 & y_1-y_0 \\ l & m \end{vmatrix}^2}}{\sqrt{l^2+m^2+n^2}}.$$

§ 13. Position relative d'une droite et d'un plan

1. Conditions de parallélisme et de perpendicularité. Soient donnés une droite L

$$\frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n}$$

et un plan π

$$Ax + By + Cz + D = 0.$$

La droite L est parallèle au plan π si et seulement si son vecteur directeur $\vec{a} = \{l;\ m;\ n\}$ est orthogonal au vecteur normal $\vec{N} = \{A;\ B;\ C\}$ du plan, c'est-à-dire si et seulement si

$$Al + Bm + Cn = 0.$$

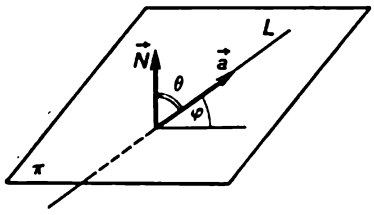


Fig. 155

La droite L est perpendiculaire au plan π si et seulement si son vecteur directeur $\vec{a}$ est colinéaire au vecteur normal $\vec{N}$ du plan π, c'est-à-dire si et seulement si

$$\frac{A}{l} = \frac{B}{m} = \frac{C}{n}.$$

2. Angle d'une droite et d'un plan. Soient donnés un plan π

$$Ax + By + Cz + D = 0$$

et une droite L

$$\frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n}$$

non perpendiculaire à π. Par angle φ de la droite L et du plan π on comprendra l'angle aigu de L et de sa projetée sur π (fig. 155). Soit θ l'angle des vecteurs $\vec{N} = \{A;\ B;\ C\}$ et $\vec{a} = \{l;\ m;\ n\}$. Si $\theta \leqslant \pi/2$ (cas de la figure 155), alors $\varphi = \pi/2 - \theta$ et $\sin\varphi = \sin(\pi/2 - \theta) = \cos\theta$. Si $\theta > \pi/2$, alors $\varphi = \theta - \pi/2$ et $\sin\varphi = \sin(\theta - \pi/2) = -\cos\theta$. Dans tous les cas $\sin\theta = |\cos\theta|$. Mais on sait calculer $\cos\theta$ [cf. § 6, formule (4)], donc

$$\sin\varphi = \frac{|Al+Bm+Cn|}{\sqrt{A^2+B^2+C^2}\sqrt{l^2+m^2+n^2}}.$$

§ 14. Quadriques

On appelle *surface du second ordre* ou *quadrique* une surface définie dans un système de coordonnées rectangulaires par une équation du second degré.

Etudions les quadriques par la méthode des sections parallèles.

1. Ellipsoïde. On appelle *ellipsoïde* la surface définie dans un système de coordonnées rectangulaires par l'équation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1. \tag{1}$$

L'équation (1) s'appelle *équation canonique de l'ellipsoïde.*

Etablissons la forme de l'ellipsoïde. Considérons à cet effet les sections de cet ellipsoïde par des plans parallèles au plan Oxy. Chacun de ces plans est défini par une équation de la forme $z = h$, où h est un nombre quelconque, et les sections, par des équations de la forme

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1-\frac{h^2}{c^2}. \tag{2}$$

Etudions l'équation (2) pour diverses valeurs de h.

1) Si $|h|>c$ $(c>0)$, alors $\frac{x^2}{a^2}+\frac{y^2}{b^2}<0$ et les sections sont des ellipses imaginaires, c'est-à-dire que les plans $z = h$ ne coupent pas l'ellipsoïde.

2) Si $h=\pm c$, alors $\frac{x^2}{a^2}+\frac{y^2}{b^2}=0$ et les sections (2) dégénèrent en deux points: $(0;0;+c)$ et $(0;0;-c)$ (les plans $z=\pm c$ sont tangents à l'ellipsoïde).

3) Si $|h|<c$, l'équation (2) peut être écrite sous la forme

$$\frac{x^2}{a^{*2}}+\frac{y^2}{b^{*2}}=1,$$

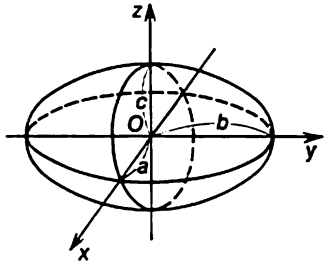


Fig. 156

les sections $z = h$ sont des ellipses homothétiques de demi-axes $a^*=a\sqrt{1-h^2/c^2}$ et $b^*=b\sqrt{1-h^2/c^2}$. Les valeurs a^* et b^* sont maximales pour $h = 0$, c'est-à-dire que la section de l'ellipsoïde par le plan de coordonnées Oxy est la plus grande ellipse.

On obtient le même tableau en envisageant les sections de l'ellipsoïde par des plans parallèles aux plans de coordonnées Oxy et Oyz.

Les sections considérées nous permettent de représenter l'ellipsoïde sous forme d'une surface ovale fermée (fig. 156). Les valeurs

a, b et c s'appellent demi-axes de l'ellipsoïde. Si $a = b = c$, l'ellipsoïde se transforme en une sphère.

2. Hyperboloïde à une nappe. On appelle *hyperboloïde à une nappe* une surface définie dans un système de coordonnées rectangulaires par l'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1. \qquad (3)$$

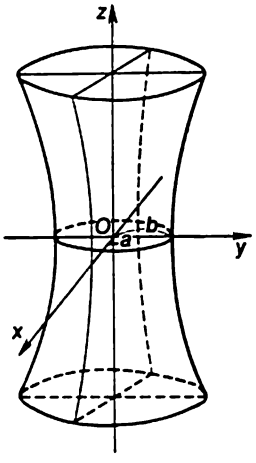


Fig. 157

L'équation (3) s'appelle *équation canonique de l'hyperboloïde à une nappe.*

Etablissons la forme de la surface (3). Considérons les sections de cette surface par les plans de coordonnées Oxz ($y = 0$) et Oyz ($x = 0$). On obtient les équations

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 \text{ et } \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

qui indiquent que ces sections sont des hyperboles.

Considérons maintenant les sections de cet hyperboloïde par les plans $z = h$. Ces sections ont pour équations

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + \frac{h^2}{c^2} \text{ ou } \frac{x^2}{a^{*2}} + \frac{y^2}{b^{*2}} = 1. \qquad (4)$$

On reconnaît ici des ellipses de demi-axes $a^* = a\sqrt{1 + h^2/c^2}$ et $b^* = b\sqrt{1 + h^2/c^2}$. La section de l'ellipsoïde par le plan Oxy est la plus petite ellipse ($a^* = a$ et $b^* = b$). Les demi-axes a^* et b^* croissent indéfiniment avec $|h|$.

On peut donc se représenter l'hyperboloïde à une nappe comme un tube qui s'élargit indéfiniment à mesure que l'on s'éloigne (d'un côté et de l'autre) du plan Oxy (fig. 157).

Les quantités a, b et c s'appellent *demi-axes de l'hyperboloïde à une nappe.* Les deux premiers demi-axes sont représentés sur la figure 157. Pour tracer le demi-axe c il faut construire le rectangle principal d'une quelconque hyperbole.

3. Hyperboloïde à deux nappes. On appelle *hyperboloïde à deux nappes* une surface définie dans un système de coordonnées rectangulaires par l'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1. \qquad (5)$$

L'équation (5) s'appelle *équation canonique de l'hyperboloïde à deux nappes.*

Etablissons la forme de la surface (5). Considérons à cet effet les sections de cette surface par les plans de coordonnées *Oxz* et *Oyz*. Ces sections ont pour équations

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = -1 \text{ et } \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1.$$

Ce sont des hyperboles.

Considérons maintenant les sections de l'hyperboloïde (5) par les plans $z = h$. Les sections $z = h$ ont pour équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{h^2}{c^2} - 1 \text{ ou}$$

$$\frac{x^2}{a^{*2}} + \frac{y^2}{b^{*2}} = 1. \qquad (6)$$

Pour $|h| > c$ $(c > 0)$ les sections $z = h$ sont des ellipses homothétiques de demi-axes $a^* = a\sqrt{h^2/c^2 - 1}$ et $b^* = b\sqrt{h^2/c^2 - 1}$. Les valeurs a^* et b^* croissent avec $|h|$.

Pour $h = \pm c$, les équations (6) ne sont vérifiées que par les coordonnées de deux points: $(0; 0; +c)$ et $(0; 0; -c)$ (les plans $z = \pm c$ sont tangents à la surface (5)).

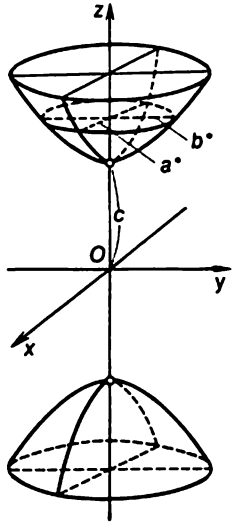


Fig. 158

Pour $|h| < c$, les sections $z = h$ sont des ellipses imaginaires, c'est-à-dire que les plans $z = h$ ne rencontrent pas l'hyperboloïde.

On peut donc se représenter l'hyperboloïde à deux nappes comme une surface composée de deux nappes (d'où son nom) ayant chacune la forme d'une cuvette infinie (fig. 158).

Les quantités a, b et c s'appellent *demi-axes de l'hyperboloïde à deux nappes.* Seule c est représentée sur la figure 158. Pour tracer a et b, il faut construire les rectangles principaux des hyperboles dans les plans *Oxz* et *Oyz*.

4. Paraboloïde elliptique. On appelle *paraboloïde elliptique* une surface définie dans un système de coordonnées rectangulaires par l'équation

$$\frac{x^2}{p} + \frac{y^2}{q} = 2z, \qquad (7)$$

où $p > 0$, $q > 0$.

L'équation (7) s'appelle *équation canonique du paraboloïde elliptique.*

Etudions cette surface par ses sections. Les sections par les plans de coordonnées Oxz et Oyz qui ont respectivement pour équations

$$x^2 = 2pz \quad \text{et} \quad y^2 = 2qz,$$

sont des paraboles de sommet O symétrique par rapport à Oz.

Considérons maintenant les sections de ce paraboloïde par les plans $z = h$. Elles ont pour équation

$$\frac{x^2}{p} + \frac{y^2}{q} = 2h \text{ ou } \frac{x^2}{a^{*2}} + \frac{y^2}{b^{*2}} = 1. \tag{8}$$

Pour $h > 0$, ce sont des ellipses homothétiques de demi-axes $a^* = \sqrt{2hp}$ et $b^* = \sqrt{2hq}$. Les valeurs a^* et b^* croissent avec h. Pour $h = 0$, la section dégénère en un point (le plan $z = 0$ est tangent au paraboloïde). Pour $h < 0$, les sections sont des ellipses imaginaires, c'est-à-dire que les plans $z = h$ ne coupent pas le paraboloïde.

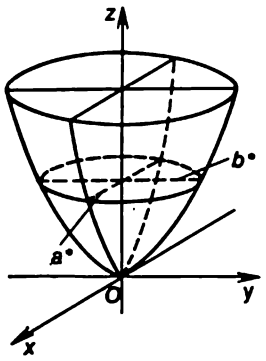


Fig. 159

Le paraboloïde elliptique a donc la forme d'une cuvette infinie (fig. 159).

Le point $(0\,;\,0\,;\,0)$ est le *sommet*, les nombres p et q, les *paramètres* du paraboloïde elliptique.

Si $p = q$, les équations (8) représentent un cercle de centre sur Oz, c'est-à-dire que le paraboloïde elliptique peut être considéré comme une surface engendrée par une parabole tournant autour de son axe. Une telle surface s'appelle *paraboloïde de révolution.*

5. Paraboloïde hyperbolique. On appelle *paraboloïde hyperbolique* une surface définie dans un système de coordonnées rectangulaires par l'équation

$$\frac{x^2}{p} - \frac{y^2}{q} = 2z, \tag{9}$$

où $p > 0$, $q > 0$.

L'équation (9) s'appelle *équation canonique du paraboloïde hyperbolique.*

Etablissons la forme géométrique de la surface (9). La section de la surface (9) par le plan Oxz a pour équations

$$\begin{cases} x^2 = 2pz \\ y = 0. \end{cases} \tag{10}$$

On reconnaît une parabole d'axe de symétrie Oz, de sommet O, tournée vers le haut. Les sections $y = h$ sont aussi les paraboles tournées vers le haut

$$x^2 = 2p\left(z + \frac{h^2}{2q}\right).$$

La section $x = 0$ est la parabole

$$y^2 = -2qz$$

d'axe de symétrie Oz, de sommet O, tournée vers le bas.

Les sections $x = h$ sont les paraboles

$$y^2 = -2q\left(z - \frac{h^2}{2p}\right)$$

tournées vers le bas dont les sommets sont situés sur la parabole définie par les équations (10).

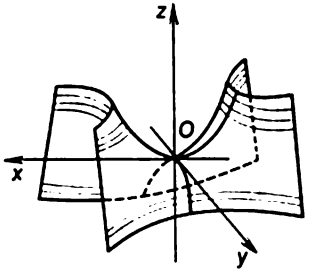


Fig. 160

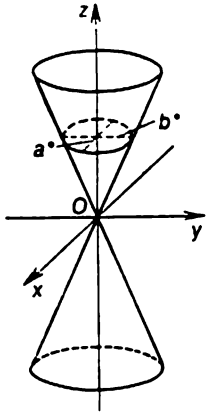


Fig. 161

Enfin les sections $z = h$ qui ont pour équation

$$\frac{x^2}{p} - \frac{y^2}{q} = 2h \text{ ou } \frac{x^2}{2ph} - \frac{y^2}{2qh} = 1,$$

sont des hyperboles coupant le plan Oxz pour $h > 0$, le plan Oyz pour $h < 0$. La section $z = 0$ est composée de deux droites concourantes

$$\frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = 0 \text{ et } \frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} = 0.$$

Le paraboloïde hyperbolique a la forme d'une selle (fig. 160). On a représenté sur cette figure les sections $z = h$ pour $h > 0$ et $h < 0$.

Le point (0 ; 0 ; 0) s'appelle *sommet*, les nombres p et q, *paramètres* du paraboloïde hyperbolique.

6. Cône du second ordre. On appelle *cône du second ordre* une surface définie dans un système de coordonnées rectangulaires par l'équation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0. \tag{11}$$

L'équation (11) s'appelle *équation canonique du cône du second ordre.*

Considérons les propriétés géométriques du cône. La section $y = 0$ est composée de deux droites

$$\frac{x^2}{a^2}-\frac{z^2}{c^2}=0 \text{ ou } \frac{x}{a}+\frac{z}{c}=0 \text{ et } \frac{x}{a}-\frac{z}{c}=0.$$

De façon analogue, la section $x = 0$ est composée de deux droites concourantes

$$\frac{y}{b}+\frac{z}{c}=0, \quad \frac{y}{b}-\frac{z}{c}=0.$$

Les sections $z = h$ sont pour $h > 0$ et $h < 0$ les ellipses

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{h^2}{c^2} \text{ ou } \frac{x^2}{a^{*2}}+\frac{y^2}{b^{*2}}=1$$

de demi-axes $a^*=\frac{a|h|}{c}$ et $b^*=\frac{b|h|}{c}$. Pour $h=0$ la section dégénère en $O(0; 0; 0)$.

Le cône du second ordre est représenté sur la figure 161.

CHAPITRE X

ÉLÉMENTS D'ALGÈBRE SUPÉRIEURE

§ 1. Matrices

1. Définition d'une matrice. Soient donnés un système de coordonnées rectangulaires $Oxyz$ et un point $M(x_1; x_2; x_3)$ *). Considérons le rayon vecteur **) $\overrightarrow{OM}$ du point M. Désignons par $\vec{e}_1$, $\vec{e}_2$ et $\vec{e}_3$ les vecteurs de base unitaires. Le vecteur $\overrightarrow{OM}$ se décompose alors de la façon suivante:

$$\overrightarrow{OM} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3.$$

Une rotation d'angle α autour de O transforme le système de coordonnées $Oxyz$ en un système $Ox'y'z'$. Désignons par x'_1, x'_2 et x'_3 les coordonnées de M par rapport à $Ox'y'z'$ et par $\vec{e}'_1$, $\vec{e}'_2$ et $\vec{e}'_3$ les vecteurs unitaires des nouveaux axes. Le vecteur $\overrightarrow{OM}$ s'écrit par rapport au nouveau système:

$$\overrightarrow{OM} = x'_1\vec{e}'_1 + x'_2\vec{e}'_2 + x'_3\vec{e}'_3.$$

En égalant les expressions de $\overrightarrow{OM}$, on obtient

$$x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 = x'_1\vec{e}'_1 + x'_2\vec{e}'_2 + x'_3\vec{e}'_3. \tag{1}$$

Voyons comment les coordonnées de M se sont transformées par la rotation du système de coordonnées.

Décomposons les vecteurs $\vec{e}'_1$, $\vec{e}'_2$ et $\vec{e}'_3$ par rapport à l'ancienne base:

$$\begin{cases} \vec{e}'_1 = a_{11}\vec{e}_1 + a_{21}\vec{e}_2 + a_{31}\vec{e}_3, \\ \vec{e}'_2 = a_{12}\vec{e}_1 + a_{22}\vec{e}_2 + a_{32}\vec{e}_3, \\ \vec{e}'_3 = a_{13}\vec{e}_1 + a_{23}\vec{e}_2 + a_{33}\vec{e}_3. \end{cases} \tag{2}$$

*) Dans ce paragraphe on désignera les coordonnées de M par x_1, x_2 et x_3 au lieu de x, y, z.

**) On appellera rayon vecteur de M le vecteur $\overrightarrow{OM}$.

Les vecteurs $\vec{e}'_1$, $\vec{e}'_2$ et $\vec{e}'_3$ étant unitaires, leurs composantes sont les cosinus directeurs :

$$a_{11} = \cos(\widehat{\vec{e}_1, \vec{e}'_1}), \quad a_{12} = \cos(\widehat{\vec{e}_1, \vec{e}'_2}), \quad a_{13} = \cos(\widehat{\vec{e}_1, \vec{e}'_3}),$$

$$a_{21} = \cos(\widehat{\vec{e}_2, \vec{e}'_1}), \quad a_{22} = \cos(\widehat{\vec{e}_2, \vec{e}'_2}), \quad a_{23} = \cos(\widehat{\vec{e}_2, \vec{e}'_3}),$$

$$a_{31} = \cos(\widehat{\vec{e}_3, \vec{e}'_1}), \quad a_{32} = \cos(\widehat{\vec{e}_3, \vec{e}'_2}), \quad a_{33} = \cos(\widehat{\vec{e}_3, \vec{e}'_3}).$$

En substituant dans (1) les décompositions (2) des vecteurs $\vec{e}'_1$, $\vec{e}'_2$ et $\vec{e}'_3$ et en regroupant les termes semblables, on obtient

$$x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 = (a_{11}x'_1 + a_{12}x'_2 + a_{13}x'_3)\vec{e}_1 + \\ + (a_{21}x'_1 + a_{22}x'_2 + a_{23}x'_3)\vec{e}_2 + (a_{31}x'_1 + a_{32}x'_2 + a_{33}x'_3)\vec{e}_3.$$

D'où

$$\begin{cases} x_1 = a_{11}x'_1 + a_{12}x'_2 + a_{13}x'_3, \\ x_2 = a_{21}x'_1 + a_{22}x'_2 + a_{23}x'_3, \\ x_3 = a_{31}x'_1 + a_{32}x'_2 + a_{33}x'_3. \end{cases} \tag{3}$$

Les coordonnées x_1, x_2 et x_3 sont donc des combinaisons linéaires des coordonnées x'_1, x'_2 et x'_3, qui sont complètement définies par les coefficients $a_{11}, \ldots, a_{33}$. Nous sommes ainsi conduits à la notion de matrice.

Définition. On appelle *matrice de la transformation* (3) le tableau

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \tag{4}$$

Cette matrice s'écrit plus brièvement

$$A = (a_{ij}) \quad (i = 1, 2, 3;\ j = 1, 2, 3);$$

a_{ij} sont les *éléments*, ou *coefficients*, de la matrice A.

La matrice A est composée de *lignes* et de *colonnes*. L'indice i désigne le numéro de la ligne, l'indice j, celui de la colonne. L'élément a_{ij} se trouve à l'intersection de la ligne i et de la colonne j.

En algèbre supérieure on étudie des matrices composées d'un nombre quelconque de lignes et de colonnes. Une telle matrice

est notée

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \qquad (5)$$

Si le nombre de lignes est égal à celui des colonnes ($m = n$), la matrice est carrée d'ordre n, sinon, rectangulaire. Ainsi la matrice (4) est une matrice carrée d'ordre trois. La matrice (5) est à m lignes et n colonnes. Si $m = 1$ et $n > 1$ on obtient une *matrice ligne* $(a_1 a_2 \dots a_n)$, si $m > 1$ et $n = 1$, une *matrice colonne*

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

Deux matrices $A = (a_{ij})$ et $B = (b_{ij})$ sont égales si leurs éléments correspondants le sont, c'est-à-dire si $a_{ij} = b_{ij}$ pour tous i et j (il est entendu que A et B ont le même nombre de lignes et de colonnes).

2. Propriétés des matrices. Comme les vecteurs, les matrices peuvent être ajoutées, multipliées entre elles et par un nombre.

1° La somme de deux matrices $A = (a_{ij})$ et $B = (b_{ij})$ de m lignes et n colonnes est la matrice $C = (c_{ij})$ d'éléments

$$a_{ij} + b_{ij} = c_{ij} \quad (i = 1, 2, \dots, m; \ j = 1, 2, \dots, n).$$

On note: $A + B = C$.

Exemple 1.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+0 & 0+0 & 0+0 \\ 0+0 & 1+0 & 0+0 \\ 0+0 & 0+0 & 0+1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

On définit de même la différence de deux matrices.

2° On appelle *produit d'une matrice* $A = (a_{ij})$ *par un scalaire* λ la matrice dont chaque élément est égal au produit de l'élément correspondant de la matrice A par λ:

$$\lambda A = \lambda (a_{ij}) = (\lambda a_{ij}) \quad (i = 1, 2, \dots, m; \ j = 1, 2, \dots, n).$$

Exemple 2.

$$3\begin{pmatrix} 1 & 0 & 2 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 & 3 \cdot 0 & 3 \cdot 2 \\ 3 \cdot 2 & 3 \cdot 0 & 3 \cdot 0 \\ 3 \cdot 0 & 3 \cdot 1 & 3 \cdot 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 6 \\ 6 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}.$$

3° On appelle *produit d'une matrice* $A = (a_{ij})$ de m lignes et k colonnes *par une matrice* $B = (b_{ij})$ de k lignes et n colonnes la matrice $C = (c_{ij})$ de m lignes et n colonnes dont l'élément c_{ij} est égal à la somme des produits des éléments de la ligne i de la matrice A par les éléments de la colonne j de la matrice B, c'est-à-dire

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ik}b_{kj}$$
$$(i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n).$$

Signalons que le nombre k de colonnes de la matrice A doit être égal au nombre des lignes de la matrice B, sinon le produit n'a pas de sens. Le produit est noté: $A \cdot B = C$.

Exemple 3.

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 0\cdot 1 + 1\cdot 0 + 0\cdot 1 & 0\cdot 0 + 1\cdot 1 + 0\cdot 0 & 0\cdot 1 + 1\cdot 0 + 0\cdot 1 \\ 1\cdot 1 + 0\cdot 0 + 1\cdot 1 & 1\cdot 0 + 0\cdot 1 + 1\cdot 0 & 1\cdot 1 + 0\cdot 0 + 1\cdot 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

Exemple 4. Soient $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, alors

$$A \cdot B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \text{ et } B \cdot A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Donc

$$A \cdot B \neq B \cdot A,$$

autrement dit, la multiplication des matrices n'est pas commutative.

Remarque. La règle de multiplication est simple à retenir sous la forme suivante: l'élément c_{ij} de C est égal au produit scalaire du vecteur ligne i de A par le vecteur colonne j de B.

On s'assure par un calcul immédiat que la somme et le produit de matrices sont justiciables des relations suivantes:

$$(A + B) \cdot C = A \cdot C + B \cdot C,$$
$$C \cdot (A + B) = C \cdot A + C \cdot B,$$
$$A \cdot (B \cdot C) = (A \cdot B) \cdot C,$$
$$(A + B) + C = A + (B + C).$$

4° *Multiplication par une matrice unité.* L'ensemble des éléments $a_{11}, a_{22}, \ldots, a_{nn}$ d'une matrice carrée $A = (a_{ij})$ s'appelle *diagonale principale de* A. On appelle *matrice unité* et on la note

E une matrice dont tous les éléments sont nuls sauf ceux de la diagonale principale qui sont égaux à 1.

Ainsi la matrice unité d'ordre 3 est

$$E=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}. \tag{6}$$

La matrice unité jouit de la remarquable propriété de ne pas modifier la matrice par laquelle elle est multipliée. C'est d'ailleurs de là qu'elle tire son nom de matrice « unité » : elle est douée de la même propriété que le nombre 1 dans la multiplication des nombres.

Exemple 5. Soient $A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$ et $E=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$. Multiplions ces matrices :

$$A\cdot E=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\cdot\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}.$$

$$E\cdot A=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\cdot\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix},$$

d'où

$$A\cdot E=A \quad \text{et} \quad E\cdot A=A.$$

Remarquons qu'à la matrice unité (6) est associée la transformation suivante des coordonnées d'un point :

$$x_1=x_1',\quad x_2=x_2',\quad x_3=x_3',$$

appelée *transformation identique*.

A la notion de matrice est étroitement liée la notion de déterminant.

§ 2. Déterminants

1. Définition d'un déterminant. Soit donnée une matrice carrée d'ordre trois

$$\begin{pmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{pmatrix}. \tag{1}$$

Définition. On appelle *déterminant d'ordre trois* de la matrice (1) le nombre désigné par le symbole

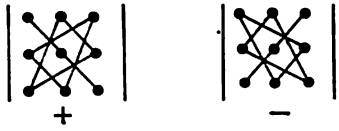
$$\Delta=\begin{vmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{vmatrix}$$

et défini par l'égalité

$$\Delta = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - c_1b_2a_3 - b_1a_2c_3 - a_1c_2b_3. \quad (2)$$

Les nombres a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2 et c_3 sont les éléments de Δ. Les éléments a_1, b_2, c_3 forment la *diagonale principale*, les éléments a_3, b_2, c_1, la *diagonale non principale*.

La *règle de Sarrus* qui se schématise comme suit



permet de retrouver facilement l'expression (2): les termes de (2) affectés du signe « + » sont égaux au produit des éléments se trouvant sur la diagonale principale et sur les sommets de deux triangles ayant un côté parallèle à cette diagonale principale. Les termes affectés du signe « — » sont égaux au produit des éléments se trouvant sur la diagonale non principale et sur les sommets de deux triangles ayant un côté parallèle à cette diagonale non principale.

Calculons par exemple

$$\begin{vmatrix} 3 & -2 & 1 \\ -2 & 1 & 3 \\ 2 & 0 & -2 \end{vmatrix} = 3\cdot1\cdot(-2)+(-2)\cdot3\cdot2+(-2)\cdot0\cdot1-$$

$$-2\cdot1\cdot1-3\cdot0\cdot3-(-2)\cdot(-2)\cdot(-2)=-12.$$

2. Propriétés des déterminants. Formulons et prouvons ces propriétés pour les déterminants d'ordre trois. A noter que ces propriétés sont valables pour les déterminants de tout ordre.

1° Un déterminant ne change pas de valeur si l'on échange ses lignes et ses colonnes, c'est-à-dire que

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Pour prouver cette propriété il suffit de calculer ces deux déterminants à l'aide de la formule (2) et de s'assurer qu'ils sont égaux.

La propriété 1° confère le même statut aux lignes et aux colonnes du déterminant. Donc les propriétés ultérieures seront formulées pour les lignes et les colonnes mais prouvées uniquement pour les unes ou les autres.

2° Un déterminant change de signe si l'on permute deux de ses lignes ou de ses colonnes. Exemple:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = - \begin{vmatrix} c_1 & b_1 & a_1 \\ c_2 & b_2 & a_2 \\ c_3 & b_3 & a_3 \end{vmatrix}.$$

Cette propriété se prouve comme la précédente.

3° Si un déterminant a deux lignes identiques ou deux colonnes identiques, il est nul.

En effet, si l'on permute les deux lignes identiques, on a d'après la propriété 2°: $\Delta = -\Delta$, d'où $\Delta = 0$. Exemple:

$$\begin{vmatrix} 1 & 1 & 17 \\ 2 & 2 & 15 \\ 3 & 3 & 9 \end{vmatrix} = 0.$$

4° Si l'on multiplie tous les éléments d'une ligne ou d'une colonne par un nombre λ, le déterminant est multiplié par λ. Exemple:

$$\begin{vmatrix} \lambda a_1 & b_1 & c_1 \\ \lambda a_2 & b_2 & c_2 \\ \lambda a_3 & b_3 & c_3 \end{vmatrix} = \lambda \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Pour prouver cette propriété il suffit de remarquer que chaque terme de la formule (2) contient un élément de chaque ligne et de chaque colonne.

5° Si tous les éléments d'une ligne ou d'une colonne sont nuls, le déterminant est nul.

Cette propriété est une conséquence de la précédente (pour $\lambda = 0$).

6° Si les éléments de deux lignes ou de deux colonnes sont proportionnels, le déterminant est nul.

En effet, si les éléments de deux colonnes sont proportionnels, en mettant le coefficient de proportionnalité en facteur on obtient un déterminant qui est nul, puisque deux de ses colonnes sont identiques (propriété 3°). Exemple:

$$\begin{vmatrix} 8 & 4 & 7 \\ 6 & 3 & 9 \\ 4 & 2 & 11 \end{vmatrix} = 2 \begin{vmatrix} 4 & 4 & 7 \\ 3 & 3 & 9 \\ 2 & 2 & 11 \end{vmatrix}.$$

7° Si chaque élément de la colonne n (resp. de la ligne n) est la somme de deux termes, le déterminant peut être représenté par la somme de deux déterminants dont la n-ième colonne (resp. ligne) de l'un est composée des premiers termes de la som-

me des éléments et la n-ième colonne (resp. ligne) de l'autre, des seconds termes de la somme des éléments, les autres colonnes (resp. lignes) étant les mêmes pour les trois déterminants. Exemple :

$$\begin{vmatrix} a_1'+a_1'' & b_1 & c_1 \\ a_2'+a_2'' & b_2 & c_2 \\ a_3'+a_3'' & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1' & b_1 & c_1 \\ a_2' & b_2 & c_2 \\ a_3' & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1'' & b_1 & c_1 \\ a_2'' & b_2 & c_2 \\ a_3'' & b_3 & c_3 \end{vmatrix}.$$

Cette propriété se prouve par un calcul direct.

8° Un déterminant ne change pas de valeur si on ajoute aux éléments d'une colonne (resp. ligne) les éléments d'une autre colonne (resp. ligne) multipliés par un nombre λ.

En effet, le déterminant obtenu peut être représenté d'après la propriété 7° par la somme de deux déterminants, l'un confondu avec le déterminant initial, l'autre nul, puisque deux de ses colonnes sont identiques. Exemple :

$$\begin{vmatrix} a_1+\lambda b_1 & b_1 & c_1 \\ a_2+\lambda b_2 & b_2 & c_2 \\ a_3+\lambda b_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} \lambda b_1 & b_1 & c_1 \\ \lambda b_2 & b_2 & c_2 \\ \lambda b_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Pour formuler la propriété suivante d'un déterminant, nous aurons besoin des notions de cofacteur et de mineur.

On appelle *mineur* d'un élément d'un déterminant le déterminant obtenu en supprimant la ligne et la colonne contenant cet élément.

Exemple : le mineur de l'élément a_1 du déterminant Δ est le déterminant d'ordre deux $\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}$, le mineur de l'élément b_1, le déterminant d'ordre deux $\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}$, etc.

On appelle *cofacteur* d'un élément d'un déterminant le mineur de cet élément multiplié par $(-1)^p$, où p est la somme des rangs de la ligne et de la colonne contenant cet élément. Le cofacteur d'un élément sera désigné par la même lettre majuscule. Ainsi le cofacteur de a_1 sera A_1, celui de b_1, B_1, etc.

Si, par exemple, l'élément a_2 appartient à la première ligne et la deuxième colonne, alors $p = 1 + 2 = 3$ et son cofacteur A_2 est

$$A_2 = (-1)^3 \cdot \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} = b_3c_1 - b_1c_3.$$

Donc le cofacteur et le mineur d'un élément ne diffèrent que par leur signe.

9° Un déterminant est égal à la somme des produits des éléments d'une ligne ou d'une colonne par leurs cofacteurs.

Autrement dit, on a les égalités suivantes:

$$\Delta = a_1A_1 + a_2A_2 + a_3A_3, \quad \Delta = a_1A_1 + b_1B_1 + c_1C_1; \tag{3}$$

$$\Delta = b_1B_1 + b_2B_2 + b_3B_3, \quad \Delta = a_2A_2 + b_2B_2 + c_2C_2; \tag{4}$$

$$\Delta = c_1C_1 + c_2C_2 + c_3C_3, \quad \Delta = a_3A_3 + b_3B_3 + c_3C_3. \tag{5}$$

Pour établir la première égalité par exemple, il suffit de mettre la formule (2) sous la forme

$$\Delta = a_1 (b_2c_3 - b_3c_2) + a_2 (b_3c_1 - b_1c_3) + a_3 (b_1c_2 - b_2c_1).$$

Les expressions entre parenthèses sont les cofacteurs des éléments a_1, a_2 et a_3, c'est-à-dire que

$$b_2c_3 - b_3c_2 = A_1, \; b_3c_1 - b_1c_3 = A_2, \; b_1c_2 - b_2c_1 = A_3.$$

D'où

$$\Delta = a_1A_1 + a_2A_2 + a_3A_3.$$

Les autres égalités (3) à (5) se prouvent de façon analogue.

La représentation d'un déterminant par l'une des formules (3) à (5) s'appelle *développement* de ce déterminant suivant les éléments d'une colonne ou d'une ligne (la première formule est le développement suivant les éléments de la première colonne).

Exemple. Calculer le déterminant

$$\Delta = \begin{vmatrix} 2 & 4 & 6 \\ 5 & 12 & 19 \\ 3 & 9 & 17 \end{vmatrix}$$

en le développant suivant les éléments de la première ligne.

Solution. On a

$$\Delta = 2\begin{vmatrix} 12 & 19 \\ 9 & 17 \end{vmatrix} - 4\begin{vmatrix} 5 & 19 \\ 3 & 17 \end{vmatrix} + 6\begin{vmatrix} 5 & 12 \\ 3 & 9 \end{vmatrix} = 8.$$

10° La somme des produits des éléments d'une colonne ou d'une ligne d'un déterminant par les cofacteurs des éléments d'une autre colonne ou d'une autre ligne est nulle.

Montrons par exemple que la somme des produits des éléments de la deuxième colonne par les cofacteurs des éléments de la première colonne est nulle. Développons à cet effet le déterminant (1) suivant les éléments de la première colonne

$$\Delta = a_1A_1 + a_2A_2 + a_3A_3. \tag{6}$$

Les cofacteurs A_1, A_2 et A_3 ne dépendent pas des éléments a_1, a_2 et a_3. Si donc l'on remplace a_1, a_2 et a_3 respectivement par h_1, h_2 et h_3 dans (6), on obtient l'égalité

$$\begin{vmatrix} h_1 & b_1 & c_1 \\ h_2 & b_2 & c_2 \\ h_3 & b_3 & c_3 \end{vmatrix} = h_1A_1 + h_2A_2 + h_3A_3. \qquad (7)$$

Si maintenant dans (7) on prend pour h_1, h_2 et h_3 les éléments b_1, b_2 et b_3 de la deuxième colonne, le déterminant du premier membre sera nul, puisque deux de ses colonnes sont identiques, et l'on obtient

$$b_1A_1 + b_2A_2 + b_3A_3 = 0.$$

C.Q.F.D.

On établit de façon analogue les égalités

$$c_1A_1 + c_2A_2 + c_3A_3 = 0,$$

$$a_1B_1 + a_2B_2 + a_3B_3 = 0, \quad a_1C_1 + a_2C_2 + a_3C_3 = 0,$$

$$c_1B_1 + c_2B_2 + c_3B_3 = 0, \quad b_1C_1 + b_2C_2 + b_3C_3 = 0$$

ainsi que les égalités relatives aux éléments des lignes:

$$a_1A_3 + b_1B_3 + c_1C_3 = 0, \quad a_2A_3 + b_2B_3 + c_2C_3 = 0,$$

$$a_1A_2 + b_1B_2 + c_1C_2 = 0, \quad a_3A_2 + b_3B_2 + c_3C_2 = 0,$$

$$a_2A_1 + b_2B_1 + c_2C_1 = 0, \quad a_3A_1 + b_3B_1 + c_3C_1 = 0.$$

§ 3. Système de trois équations du premier degré à trois inconnues

La théorie des matrices et des déterminants est largement utilisée aussi bien en mathématiques que dans ses applications. C'est un outil d'investigation dont le spectre est immense.

Appliquons les matrices et les déterminants à l'étude d'un système de trois équations du premier degré à trois inconnues x, y, z:

$$\begin{cases} a_1x + b_1y + c_1z = h_1, \\ a_2x + b_2y + c_2z = h_2, \\ a_3x + b_3y + c_3z = h_3 \end{cases} \qquad (1)$$

(les coefficients a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3 et les termes constants h_1, h_2 et h_3 sont donnés).

On dit qu'un triplet de nombres x_0, y_0, z_0 est *solution du système* (1) si la substitution de ces nombres aux inconnues x, y, z transforme les trois équations (1) en identités.

Les déterminants suivants sont essentiels pour la suite :

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \quad \Delta_x = \begin{vmatrix} h_1 & b_1 & c_1 \\ h_2 & b_2 & c_2 \\ h_3 & b_3 & c_3 \end{vmatrix},$$

$$\Delta_y = \begin{vmatrix} a_1 & h_1 & c_1 \\ a_2 & h_2 & c_2 \\ a_3 & h_3 & c_3 \end{vmatrix}, \quad \Delta_z = \begin{vmatrix} a_1 & b_1 & h_1 \\ a_2 & b_2 & h_2 \\ a_3 & b_3 & h_3 \end{vmatrix}.$$

Le déterminant Δ s'appelle *déterminant principal* du système (1), les déterminants Δ_x, Δ_y et Δ_z obtenus en remplaçant respectivement la première, la deuxième et la troisième colonne par la colonne des termes constants, *déterminants caractéristiques.*

Traitons les cas où $\Delta \neq 0$ et $\Delta = 0$.

Cas 1. $\Delta \neq 0$. Montrons que le système (1) admet une solution et une seule. A cet effet multiplions les deux membres de la première, de la deuxième et de la troisième équation du système (1) respectivement par les cofacteurs A_1, A_2 et A_3 et additionnons. On obtient

$$(a_1A_1 + a_2A_2 + a_3A_3)\,x + (b_1A_1 + b_2A_3 + b_3A_3)\,y + \\ + (c_1A_1 + c_2A_2 + c_3A_3)\,z = h_1A_1 + h_2A_3 + h_3A_3.$$

Les propriétés 9° et 10° des déterminants nous donnent

$$\Delta \cdot x = h_1A_1 + h_2A_2 + h_3A_3 = \Delta_x.$$

On trouve de façon analogue

$$\Delta \cdot y = h_1B_1 + h_2B_2 + h_3B_3 = \Delta_y,$$

$$\Delta \cdot z = h_1C_1 + h_2C_2 + h_3C_3 = \Delta_z.$$

Le système (1) nous conduit donc au système d'équations

$$x = \frac{\Delta_x}{\Delta}, \quad y = \frac{\Delta_y}{\Delta}, \quad z = \frac{\Delta_z}{\Delta}. \tag{2}$$

Les formules (2) s'appellent *formules de Cramer* *).

Pour montrer que le système (1) admet une solution, substituons à x, y, z les valeurs définies par les formules de Cramer et assurons-nous que les équations (1) se transforment toutes trois

*) Cramer Gabriel (1704-1752), mathématicien suisse.

en identités. Pour la première équation on a

$$a_1x+b_1y+c_1z=a_1\frac{\Delta_x}{\Delta}+b_1\frac{\Delta_y}{\Delta}+c_1\frac{\Delta_z}{\Delta}=$$

$$=\frac{1}{\Delta}\{a_1(h_1A_1+h_2A_2+h_3A_3)+$$

$$+b_1(h_1B_1+h_2B_2+h_3B_3)+c_1(h_1C_1+h_2C_2+h_3C_3)\}=$$

$$=\frac{1}{\Delta}\{h_1(a_1A_1+b_1B_1+c_1C_1)+h_2(a_1A_2+b_1B_2+c_1C_2)+$$

$$+h_3(a_1A_3+b_1B_3+c_1C_3)\}=h_1,$$

puisque

$$a_1A_1+b_1B_1+c_1C_1=\Delta$$

d'après la propriété 9° et

$$a_1A_2+b_1B_2+c_1C_2=0,\quad a_1A_3+b_1B_3+c_1C_3=0$$

d'après la propriété 10°.

On a ainsi établi que la première équation du système (1) se transforme en identité. On démontre de même que la deuxième et la troisième équation du système se transforment en identités.

Donc le système (1) admet une solution.

Les formules de Cramer établissent aussi l'unicité de la solution du système (1), puisque le système (2) découle du système (1) et donc toute solution de (1) est solution de (2), c'est-à-dire s'exprime par les formules de Cramer.

Ce qui précède nous permet de tirer la conclusion suivante: si le déterminant Δ du système (1) est différent de zéro, ce système admet une solution et une seule qui est exprimée par les formules de Cramer.

Exemple. Trouver les solutions du système

$$\begin{cases} x+2y+z=4, \\ 3x-5y+3z=1, \\ 2x+7y-z=8. \end{cases}$$

Solution. Puisque $\Delta=33\neq 0$, ce système admet une seule solution définie par les formules (2):

$$x=\frac{\Delta_x}{x}=\frac{33}{33}=1,\quad y=\frac{\Delta_y}{\Delta}=\frac{33}{33}=1,\quad z=\frac{\Delta_z}{\Delta}=\frac{33}{33}=1.$$

Donc $x=1$, $y=1$, $z=1$ est la solution de ce système.

Cas 2. $\Delta=0$. Supposons que l'un au moins des déterminants caractéristiques Δ_x, Δ_y, Δ_z est non nul. Dans ces conditions

l'une au moins des égalités (2) est impossible *), c'est-à-dire que le système (2) n'admet pas de solution, donc le système (1) n'en admet pas non plus puisque le système (2) découle du système (1).

Supposons maintenant que les déterminants Δ_x, Δ_y et Δ_z sont nuls. Etudions tout d'abord les systèmes homogènes.

On appelle *système homogène* de trois équations du premier degré à trois inconnues un système de la forme

$$\begin{cases} a_1x + b_1y + c_1z = 0, \\ a_2x + b_2y + c_2z = 0, \\ a_3x + b_3y + c_3z = 0. \end{cases} \tag{3}$$

Il est évident que ce système admet toujours la solution triviale $x = 0$, $y = 0$, $z = 0$. Si $\Delta \neq 0$, cette solution est unique (en vertu du cas 1).

Montrons que si le déterminant $\Delta = 0$, le système (3) admet une infinité de solutions non nulles. Faisons la démonstration en deux étapes.

1) Supposons que l'un au moins des mineurs du déterminant Δ est non nul, pour fixer les idées $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$. Le système homogène composé des deux premières équations (3) peut être mis sous la forme

$$\begin{cases} a_1x + b_1y = -c_1z, \\ a_2x + b_2y = -c_2z. \end{cases} \tag{4}$$

Considérons les cofacteurs A_3, B_3, C_3 des éléments de la troisième ligne du déterminant :

$$A_3 = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \quad B_3 = -\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}, \quad C_3 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Puisque $C_3 \neq 0$ par hypothèse, le système (4) admet une seule solution pour chaque z. Cette solution peut être mise sous la forme $x = \frac{A_3}{C_3}z$, $y = \frac{B_3}{C_3}z$. Posons $z = C_3t$, où t peut prendre toute valeur. Il est alors évident que le système homogène (4) admet une infinité de solutions définies par les formules

$$x = A_3t, \quad y = B_3t, \quad z = C_3t, \tag{5}$$

où t est un nombre arbitraire.

Reste à montrer que les solutions x, y, z définies par (5) transforment la troisième équation du système (3) en identité. En effet, en portant les expressions (5) de x, y, z dans le premier

*) En effet, supposons par exemple que $\Delta_x \neq 0$. L'égalité $\Delta \cdot x = \Delta_x$ est alors impossible, puisque $\Delta \cdot x = 0$ pour tout x, alors que $\Delta_x \neq 0$.

membre de la troisième équation, on trouve

$$a_3x + b_3y + c_3z = (a_3A_3 + b_3B_3 + c_3C_3)\,t = \Delta \cdot t = 0 \cdot t = 0.$$

Donc les formules (5) définissent une solution du système homogène (3) pour tout t.

2) Supposons maintenant que tous les mineurs du déterminant Δ sont nuls. Ceci exprime que les coefficients des équations (3) sont proportionnels. La deuxième et la troisième équation qui découlent alors de la première peuvent être éliminées et l'on obtient une seule équation à trois inconnues $a_1x + b_1y + c_1z = 0$ qui admet visiblement une infinité de solutions (on peut attribuer des valeurs arbitraires à deux de ces inconnues et résoudre l'équation par rapport à la troisième).

On a ainsi démontré que si $\Delta = 0$, le système homogène (3) admet une infinité de solutions.

Etudions maintenant le système (1) dans le cas où les déterminants Δ, Δ_x, Δ_y, Δ_z sont tous nuls. Montrons que si le système (1) admet une solution, il en admettra une infinité.

Soit x_0, y_0, z_0 une solution du système (1). On a les identités

$$\begin{cases} a_1x_0 + b_1y_0 + c_1z_0 = h_1, \\ a_2x_0 + b_2y_0 + c_2z_0 = h_2, \\ a_3x_0 + b_3y_0 + c_3z_0 = h_3. \end{cases} \tag{6}$$

En soustrayant membre à membre les identités (6) des équations (1) on obtient un système d'équations équivalent à (1):

$$\begin{cases} a_1(x - x_0) + b_1(y - y_0) + c_1(z - z_0) = 0, \\ a_2(x - x_0) + b_2(y - y_0) + c_2(z - z_0) = 0, \\ a_3(x - x_0) + b_3(y - y_0) + c_3(z - z_0) = 0. \end{cases} \tag{7}$$

C'est un système homogène de trois équations du premier degré aux inconnues $(x - x_0)$, $(y - y_0)$ et $(z - z_0)$ dont le déterminant Δ est nul. Mais d'après ce qui vient d'être prouvé, ce système admet une infinité de solutions, donc le système (1) aussi. Si, par exemple, le mineur C_3 est non nul, la solution du système (1) peut être mise sous la forme (en vertu de la formule (5))

$$x = x_0 + A_3t,\ y = y_0 + B_3t,\ z = z_0 + C_3t,$$

où t prend toute valeur. Ce qu'on voulait. Donc si $\Delta = \Delta_x = \Delta_y = \Delta_z = 0$, le système (1) ne possède pas de solutions ou bien en possède une infinité.

A titre d'exemple nous proposons au lecteur d'étudier les trois systèmes suivants:

$$\begin{cases} x+y+z=2, \\ 3x+2y+2z=1, \\ 4x+3y+3z=4, \end{cases} \quad \begin{cases} x+y+z=1, \\ 2x+y+z=2, \\ 3x+2y+2z=3, \end{cases} \quad \begin{cases} x+y+z=1, \\ 2x+2y+2z=3, \\ 3x+3y+3z=4 \end{cases}$$

et de s'assurer que le premier ne possède pas de solutions ($\Delta = 0$, $\Delta_y = 1 \neq 0$), le deuxième en possède une infinité ($\Delta = \Delta_x = \Delta_y = \Delta_z = 0$) définies par les formules $x = 1$, $y = t$, $z = -t$, le troisième n'en possède aucune ($\Delta = \Delta_x = \Delta_y = \Delta_z = 0$). Dans le troisième système, on pourra s'assurer de l'incompatibilité des deux premières équations.

§ 4. Ecriture matricielle d'un système d'équations linéaires. Notion de matrice inverse

Considérons de nouveau le système d'équations (1) du § 3:

$$\begin{cases} a_1x+b_1y+c_1z=h_1, \\ a_2x+b_2y+c_2z=h_2, \\ a_3x+b_3y+c_3z=h_3. \end{cases} \tag{1}$$

Introduisons les notations suivantes:

$$A=\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}, \quad X=\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad H=\begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}. \tag{2}$$

En utilisant la règle de multiplication des matrices, on peut mettre le système (1) sous la forme matricielle équivalente

$$AX = H, \tag{3}$$

où A est une matrice donnée; H, une matrice colonne donnée; X, une matrice colonne inconnue. Une solution de l'équation (3) est une matrice colonne X qui transforme l'équation (3) en identité.

Supposons que le déterminant Δ de la matrice A est non nul. Alors, comme établi au § 3, le système (1), donc le système (3), possèdent une solution unique donnée par les formules de Cramer. On se propose de représenter la solution de l'équation (3) sous une forme matricielle. Introduisons à cet effet la notion de matrice inverse.

On appelle *inverse* d'une matrice A une matrice notée A^{-1} telle que

$$A^{-1} \cdot A = A \cdot A^{-1} = E, \tag{4}$$

où E est la matrice unité.

Montrons que si $\Delta \neq 0$, l'inverse de la matrice A est la matrice suivante :

$$A^{-1} = \begin{pmatrix} A_1/\Delta & A_2/\Delta & A_3/\Delta \\ B_1/\Delta & B_2/\Delta & B_3/\Delta \\ C_1/\Delta & C_2/\Delta & C_3/\Delta \end{pmatrix}, \tag{5}$$

où A_i, B_i, C_i désignent les cofacteurs des éléments a_i, b_i, c_i ($i = 1, 2, 3$) respectivement. Pour cela il nous faut prouver que la matrice (5) satisfait aux conditions (4). Vérifions par exemple l'égalité

$$A \cdot A^{-1} = E.$$

En multipliant A par A^{-1}, on obtient

$$AA^{-1} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} A_1/\Delta & A_2/\Delta & A_3/\Delta \\ B_1/\Delta & B_2/\Delta & B_3/\Delta \\ C_1/\Delta & C_2/\Delta & C_3/\Delta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E.$$

On s'est servi du fait qu'en vertu des propriétés 9° et 10°

$$a_iA_j + b_iB_j + c_iC_j = \begin{cases} \Delta \text{ si } i = j, \\ 0 \text{ sinon.} \end{cases}$$

Donc, en particulier, l'élément de la matrice AA^{-1} qui est contenu dans la première ligne et la première colonne est égal à $(a_1A_1 + b_1B_1 + c_1C_1)/\Delta = \Delta/\Delta = 1$ et l'élément qui appartient à la première ligne et la deuxième colonne à $(a_1A_2 + b_1B_2 + c_1C_2)/\Delta = 0$.

On s'assure que tous les autres éléments de la matrice AA^{-1} sont égaux aux éléments correspondants de la matrice E. Ainsi $AA^{-1} = E$. On démontre de même que $A^{-1}A = E$.

Donc l'inverse de la matrice A est la matrice A^{-1} définie par la formule (5). De l'égalité (4) il s'ensuit que la matrice A est l'inverse de la matrice A^{-1}. Les matrices A et A^{-1} sont donc inverses l'une de l'autre.

Remarque. Si le déterminant $\Delta = 0$, la matrice inverse A^{-1} n'existe pas.

Utilisons la matrice inverse pour résoudre l'équation (3). Une multiplication de l'équation (3) à gauche par la matrice A^{-1} nous donne

$$A^{-1}AX = A^{-1}H. \tag{6}$$

Comme $A^{-1}A = E$ et $EX = X$, on déduit de (6)

$$X = A^{-1}H. \tag{7}$$

Il est immédiat de s'assurer que l'expression de X est bien la solution de l'équation (3). En effet, en la portant dans (3) on

obtient

$$AA^{-1}H = H.$$

Cette égalité est une identité, puisque $AA^{-1} = E$ et $EH = H$.

Si donc $\Delta \neq 0$, la solution de l'équation (3), donc du système (1), peut être mise sous la forme matricielle (7). Cette solution est la même que celle qui a été acquise au § 3 par les formules de Cramer. Ce fait qui découle de l'unicité de la solution du système (1) pour $\Delta \neq 0$ se vérifie immédiatement par la substitution, dans la formule (7), de l'expression (5) de A^{-1} et des expressions (2) de X et de H. Alors

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} A_1/\Delta & A_2/\Delta & A_3/\Delta \\ B_1/\Delta & B_2/\Delta & B_3/\Delta \\ C_1/\Delta & C_2/\Delta & C_3/\Delta \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix},$$

d'où $x = (h_1A_1 + h_2A_2 + h_3A_3)/\Delta = \Delta_x/\Delta$, $y = (h_1B_1 + h_2B_2 + h_3B_3)/\Delta = \Delta_y/\Delta$, $z = (h_1C_1 + h_2C_2 + h_3C_3)/\Delta = \Delta_z/\Delta$.

Exemple. Résoudre le système d'équations

$$\begin{cases} x + 2y + z = 1, \\ 2x + y + z = -1, \\ x + 3y + z = 2. \end{cases}$$

Solution. On a

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

La matrice A possède une matrice inverse A^{-1} puisque son déterminant est égal à 1. La formule (5) nous donne

$$A^{-1} = \begin{pmatrix} -2 & 1 & 1 \\ -1 & 0 & 1 \\ 5 & -1 & -3 \end{pmatrix}.$$

D'après la formule (7)

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 & 1 & 1 \\ -1 & 0 & 1 \\ 5 & -1 & -3 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix},$$

d'où $x = -1$, $y = 1$, $z = 0$.

Signalons en conclusion que l'étude des systèmes d'équations du premier degré à plusieurs inconnues et de nombreux problèmes de mathématiques nous conduit à envisager des matrices et des déterminants d'ordre n arbitraire ($n = 2, 3, 4, \ldots$). La

théorie des matrices et déterminants d'ordre n se construit par analogie à celle des matrices et déterminants d'ordre trois. Cependant son élaboration rigoureuse implique l'introduction d'autres notions et la démonstration de nombreux théorèmes complexes. Ceux qui voudraient élargir et approfondir leurs connaissances sur cette théorie et sur la théorie des systèmes d'équations du premier degré à plusieurs inconnues peuvent s'adresser à n'importe quel cours d'algèbre supérieure *).

*) Par exemple à A. Kurosh, *Cours d'algèbre supérieure*. Trad. française. Ed. Mir, 1984.

CHAPITRE XI

FONCTIONS DE PLUSIEURS VARIABLES : NOTION, LIMITE ET CONTINUITÉ

§ 1. Notion de fonction de plusieurs variables

1. Remarques préliminaires. Jusqu'à maintenant nous avons étudié des fonctions d'une variable.

L'étude de nombreux problèmes nous conduit à envisager des relations entre plusieurs variables dans lesquelles les valeurs de l'une d'elles sont complètement définies par celles des autres. Ainsi, par exemple, la température d'un corps à un instant t peut varier d'un point à l'autre. Chaque point de ce corps est défini par trois coordonnées x, y, z, donc la température dépend des trois variables x, y et z auxquelles peut s'ajouter une quatrième si l'on fait intervenir le temps t. L'aire d'un rectangle de côtés x, y dépend de deux variables x, y ; le volume d'un parallélépipède rectangle de dimensions x, y, z dépend de trois variables x, y, z. On peut multiplier à l'infini les exemples de telles relations.

Cette partie du cours est consacrée à l'étude de telles relations. A cet effet introduisons la notion de fonction de plusieurs variables et développons l'appareil mathématique nécessaire à l'étude de ces fonctions.

2. Définition d'une fonction de deux et de plusieurs variables. La fonction de deux variables se définit comme celle d'une variable.

Définition. Soient X, Y et Z des ensembles de nombres. On appelle *fonction de deux variables* l'ensemble f des triplets ordonnés $(x ; y ; z)$, $x \in X$, $y \in Y$, $z \in Z$, tels que chaque couple ordonné $(x ; y)$ figure dans un triplet et un seul de cet ensemble, et chaque z, dans au moins un triplet. Ceci étant, on dit qu'au couple ordonné de nombres $(x ; y)$ est associé le nombre z et on note $z = f(x ; y)$. Le nombre z s'appelle *valeur de la fonction f au point* $(x ; y)$. La variable z s'appelle *variable dépendante*, les variables x et y, *variables indépendantes* (ou *arguments*) ; l'ensemble $\{(x ; y)\}$, *ensemble* ou *domaine de définition de f*, l'ensemble Z, *ensemble* ou *domaine des valeurs de f*.

Puisque à tout couple ordonné de nombres $(x ; y)$ est associé un point unique M du plan rapporté à un système de coordonnées rectangulaires fixé et, réciproquement, à chaque point M corres-

pond un couple ordonné unique de nombres $(x; y)$, une fonction de deux variables peut être traitée comme une fonction du point M et au lieu de $z = f(x, y)$, on peut écrire $z = f(M)$. L'ensemble de définition de la fonction est dans ce cas un ensemble $\{M\}$ de points du plan. Dans la suite on utilisera les deux notations d'une fonction de deux variables.

Comme dans le cas d'une variable, il existe plusieurs procédés de définition d'une fonction de deux variables. Dans les exemples on se donnera généralement une fonction sous forme analytique, c'est-à-dire par une formule. Dans ce cas l'ensemble de définition de la fonction sera l'ensemble de tous les points du plan en lesquels cette formule a un sens.

Considérons des exemples de fonctions de deux variables.

1. $z = x^2 + y^2$. L'ensemble de définition de cette fonction est l'ensemble $\{M\}$ de tous les couples de points $(x; y)$, c'est-à-dire le plan Oxy tout entier, l'ensemble des valeurs, l'intervalle $Z = [0, +\infty[$.

2. $z = \sqrt{1 - x^2 - y^2}$. L'ensemble de définition de cette fonction est l'ensemble de tous les points en lesquels l'expression $\sqrt{1 - x^2 - y^2}$ a un sens, c'est-à-dire l'ensemble des points en lesquels $1 - x^2 - y^2 \geqslant 0$ ou $x^2 + y^2 \leqslant 1$. Cet ensemble est un disque de centre O et de rayon 1. L'ensemble des valeurs est l'intervalle $[0, 1]$.

3. $z = 1/\sqrt{x^2 + y^2 - 1}$. L'ensemble de définition de cette fonction est l'ensemble des points dont les coordonnées satisfont à l'inégalité $x^2 + y^2 > 1$, c'est-à-dire l'ensemble des points extérieurs au disque de rayon 1 centré en O. L'ensemble des valeurs est l'intervalle $]0, +\infty[$.

De ces exemples il ressort qu'une fonction de deux variables peut être définie sur tout ou partie du plan Oxy.

Du cours de géométrie analytique on sait que l'ensemble de tous les triplets ordonnés $(x; y; z)$ engendre un espace de coordonnées. A chaque triplet $(x; y; z)$ est associé un point unique $M(x; y; z)$ de cet espace et réciproquement. Si au lieu d'un ensemble $\{M\}$ de points du plan, on envisage un ensemble $\{M\}$ de points de l'espace, on peut définir par analogie une fonction $u = f(M)$ ou $u = f(x; y; z)$. Une fonction de trois variables peut être définie sur tout ou partie de l'espace. Ainsi la fonction $u = x^2 + y^2 + z^2$ est définie sur l'espace tout entier, la fonction $u = \ln xyz$, sur l'ensemble des points dont les coordonnées satisfont à l'inégalité $xyz > 0$. Les ensembles des valeurs de ces fonctions sont respectivement les intervalles $[0, +\infty[$ et $]-\infty, +\infty[$.

On définit de même une fonction de quatre variables $u = f(x, y, z, t)$. Dans ce cas l'ensemble des quadriplets ordonnés $(x; y; z; t)$ engendre un espace à quatre dimensions. Chaque

quadriplet $(x; y; z; t)$ est un point de cet espace. Cependant on ne peut plus se représenter géométriquement le domaine de définition d'une fonction de quatre variables. On définit de façon analogue les fonctions de cinq et, d'une façon plus générale, de n variables $u = f(x_1, x_2, \ldots, x_n)$.

Dans la suite on étudie en détail les fonctions de deux variables; à noter que la généralisation des définitions et des résultats obtenus aux fonctions de trois et de plusieurs variables n'apporte en principe aucune complication.

§ 2. Représentation géométrique des fonctions de deux variables

On sait qu'une fonction d'une variable se représente dans le plan par une courbe d'équation $y = f(x)$. Une fonction de deux variables se représente dans l'espace par une surface d'équation $z = f(x, y)$.

En géométrie analytique on étudie les surfaces et leurs équations. Par exemple, l'équation $z - 2x + 5y + 10 = 0$ est l'équation d'un plan. Ce plan est le graphique de la fonction $z = 2x - 5y - 10$.

L'équation $x^2 + y^2 + z^2 = R^2$ est l'équation d'une sphère de rayon R et de centre O. Par ailleurs, la sphère est la réunion des graphiques de deux fonctions: $z = \sqrt{R^2 - x^2 - y^2}$ et $z = -\sqrt{R^2 - x^2 - y^2}$.

La construction des graphiques de fonctions de deux variables est souvent très compliquée. Il existe cependant une méthode de représentation graphique des fonctions de deux variables basée sur les sections de la surface $z = f(x, y)$ par des plans $z = c$, où c est un nombre quelconque.

On appelle *ligne de niveau* d'une fonction $z = f(x, y)$ l'ensemble des points (x, y) du plan Oxy en lesquels cette fonction prend la même valeur c. Il est évident qu'on obtient des lignes de niveau différentes en faisant varier c.

Si les nombres $c_1, c_2, \ldots, c_n$ forment une progression arithmétique de raison h, la position relative des lignes de niveau correspondantes nous donne une idée générale de la forme de la surface. Si les lignes de niveau sont fortement concentrées, la fonction croît rapidement (et la pente est raide); si les lignes de niveau sont faiblement concentrées, la fonction croît lentement (et la pente est douce) (fig. 162). Il est clair que plus h est petit et plus la représentation de la fonction est précise.

La « ligne de niveau » a été empruntée à la cartographie, où elle représente les points de même altitude. La ligne de niveau donne une idée non seulement de l'altitude d'un point d'une région, mais aussi du relief, ce qui est particulièrement important si cette région est montagneuse.

Exemple. Construire les lignes de niveau de la fonction $z = x^2 + y^2$.

Solution. Les lignes de niveau de cette fonction sont définies par l'équation $x^2 + y^2 = c$ $(0 \leqslant c \leqslant +\infty)$. En faisant varier c on obtient une famille de cercles concentriques. Pour $c = 0$ la ligne de niveau dégénère en le point (0; 0) (fig. 163).

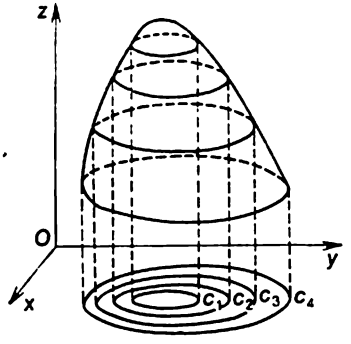


Fig. 162

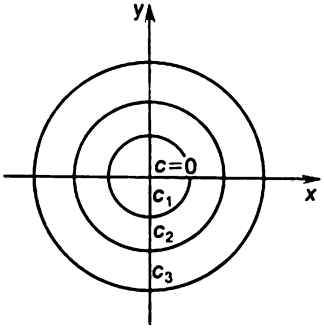


Fig. 163

Puisque les lignes de niveau sont ici des cercles centrés en O, la représentation graphique de cette fonction est une surface de révolution autour de l'axe Oz. En fait, on sait du cours de géométrie analytique que l'équation $z = x^2 + y^2$ représente un paraboloïde de révolution.

§ 3. Limite d'une fonction de deux variables

Introduisons la notion de δ-voisinage d'un point donné $M_0\,(x_0;\ y_0)$ et la notion de suite convergente de points du plan.

Définition 1. On appelle δ-voisinage d'un point $M_0\,(x_0;\ y_0)$ l'ensemble $\{M\,(x;\ y)\}$ des points dont les coordonnées x et y vérifient l'inégalité $\sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$ ou, plus brièvement, $\rho\,(M;\ M_0) < \delta$.

En d'autres termes, un δ-voisinage du point M_0 est l'ensemble des points intérieurs à un cercle de centre M_0 et de rayon δ. ou encore un disque ouvert de centre M_0 et de rayon δ.

Considérons une suite de points $M_1\,(x_1;\ y_1)$, $M_2\,(x_2;\ y_2)$, ..., $M_n\,(x_n;\ y_n)$, ... que nous désignerons brièvement par $\{M_n\}$*).

Définition 2. On dit qu'une suite de points $\{M_n\}$ *converge* vers un point M_0 si pour tout $\varepsilon > 0$ il existe un entier N tel que pour $n > N$ l'on ait $\rho\,(M_n;\ M_0) < \varepsilon$.

*) La donnée d'une suite $\{M_n\}$ équivaut à la donnée de deux suites numériques $\{x_n\}$ et $\{y_n\}$, puisqu'un point du plan est défini par deux coordonnées: x et y.

Le point M_0 s'appelle *limite de la suite* $\{M_n\}$.

On note ceci: $\lim\limits_{n\to\infty} M_n = M_0$ ou $M_n \to M_0$ lorsque $n \to \infty$.

Signalons que la notion de suite convergente de points du plan généralise celle de suite numérique convergente. En effet, la donnée d'une suite $\{M_n\}$ de points de la droite équivaut à la donnée d'une suite numérique $\{x_n\}$ et l'inégalité $\rho(M_n; M_0) < \varepsilon$ se transforme dans ce cas en l'inégalité $|x_n - x_0| < \varepsilon$.

Définissons maintenant la limite d'une fonction de deux variables. Cette définition est identique à celle de la limite d'une fonction d'une variable *).

Soit $z = f(M)$ une fonction définie sur un ensemble $\{M\}$ et soit $M_0 \in \{M\}$ ou $M_0 \notin \{M\}$ un point dont tout δ-voisinage contient au moins un point de $\{M\}$ autre que M_0.

Définition 3. On dit qu'un nombre A est la *limite de la fonction* $z = f(M)$ *au point* M_0 si pour toute suite de points $\{M_n\}$ ($M_n \neq M_0$, $M_n \in \{M\}$) la suite correspondante $\{f(M_n)\}$ des valeurs de la fonction f converge vers A.

On note ceci:

$$\lim_{M\to M_0} f(M) = A \quad \text{ou} \quad \lim_{\substack{x\to x_0\\ y\to y_0}} f(x, y) = A.$$

Trouvons, par exemple, la limite de la fonction $f(x, y) = x^2 + y^2$ au point $M_0(1; 2)$. Cette fonction est définie sur le plan tout entier. Pour toute suite de points $\{M_n\}$ convergeant vers M_0, on a

$$\lim_{n\to\infty}(x_n^2 + y_n^2) = \lim_{n\to\infty} x_n^2 + \lim_{n\to\infty} y_n^2 = 1^2 + 2^2 = 5,$$

donc $\lim\limits_{\substack{x\to 1\\ y\to 2}} (x^2 + y^2) = 5$.

Citons un exemple de fonction ne possédant pas de limite en un point. La fonction $f(x, y) = \dfrac{x-y}{x+y}$ est partout définie sauf sur la droite $x + y = 0$. Montrons qu'elle n'admet pas de limite au point $(0; 0)$. Choisissons à cet effet deux suites de points $M_n(1/n; 0)$ et $M_n'(0; 1/n)$ convergentes vers le point $(0; 0)$; alors

$$\lim_{n\to\infty} f(M_n) = \lim_{n\to\infty} \frac{1/n - 0}{1/n + 0} = 1 \quad \text{et} \quad \lim_{n\to\infty} f(M_n') = \lim_{n\to\infty} \frac{0 - 1/n}{0 + 1/n} = -1.$$

Donc à deux suites de points différentes convergeant vers l'origine des coordonnées correspondent deux suites de valeurs de la fonction possédant des limites différentes. Par conséquent, cette fonction n'admet pas de limite en $(0; 0)$ d'après la définition 3.

*) Il est recommandé de reprendre la lecture du chap. IV, § 2, n° 1.

La limite d'une fonction de deux variables a été définie à l'aide de la notion de limite d'une suite. Comme pour une fonction d'une variable, on peut produire une définition équivalente en termes de ε et δ.

Définition 4. On dit qu'un nombre A est la *limite d'une fonction* $z = f(M)$ *en un point* M_0 si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout point $M \in \{M\}$ vérifiant la condition $0 < \rho(M; M_0) < \delta$, l'on ait $|f(M) - A| < \varepsilon$.

En signes logiques, cette définition s'écrit

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall M \in \{M\},\ 0 < \rho(M; M_0) < \delta : \\ : |f(M) - A| < \varepsilon.$$

L'équivalence des définitions 3 et 4 se prouve comme pour une fonction d'une variable, sauf que dans la démonstration du théorème 4.1 il faut remplacer la suite $\{x_n\}$ par la suite $\{M_n\}$, le point x_0, par le point M_0, les différences $|x - x_0|$ et $|x_n - x_0|$ respectivement par les distances $\rho(M; M_0)$ et $\rho(M_n; M_0)$ et enfin la suite numérique $\{f(x_n)\}$ par la suite numérique $\{f(M_n)\}$.

La définition de la limite d'une fonction de deux variables nous permet d'étendre les théorèmes fondamentaux des limites d'une fonction d'une variable à une fonction de deux variables. On a par exemple le théorème suivant.

Théorème. *Si* $f(M)$ *et* $g(M)$ *sont des fonctions définies sur le même ensemble* $\{M\}$ *et possédant les limites* B *et* C *en* M_0, *les fonctions* $f(M) \pm g(M)$, $f(M) \cdot g(M)$ *et* $\frac{f(M)}{g(M)}$ $(C \neq 0)$ *admettent en* M_0 *des limites respectivement égales à* $B \pm C$, $B \cdot C$ *et* $\frac{B}{C}$.

La démonstration de ce théorème est identique à celle du théorème 4.3 : elle peut en être déduite par un changement formel des lettres x et x_0 en M et M_0 et par une utilisation de la définition 3 de la limite d'une fonction de deux variables à la place de la définition 1 de la limite d'une fonction d'une variable.

Définition 5. On dit qu'une fonction $z = f(M)$ est un *infiniment petit* en un point $M = M_0$ (ou lorsque $M \to M_0$) si $\lim\limits_{M \to M_0} f(M) = 0$.

Si la fonction $z = f(M)$ admet au point M_0 une limite égale à A, la fonction $\alpha(M) = f(M) - A$ est un infiniment petit au point M_0. En effet, $\lim\limits_{M \to M_0} \alpha(M) = \lim\limits_{M \to M_0} [f(M) - A] = \lim\limits_{M \to M_0} f(M) - \lim\limits_{M \to M_0} A = A - A = 0$. D'où l'on déduit une représentation spéciale de la fonction qui admet au point M_0 une limite égale à A : $f(M) = A + \alpha(M)$, où $\lim\limits_{M \to M_0} \alpha(M) = 0$.

Les fonctions infiniment petites de deux variables sont justiciables des mêmes règles de comparaison que les fonctions infiniment petites d'une variable. Comme dans le cas d'une variable, par $o(\beta)$ on comprendra tout infiniment petit en M_0 d'ordre supérieur à β (M), c'est-à-dire que $\lim\limits_{M \to M_0} \frac{o(\beta)}{\beta} = 0$.

§ 4. Continuité d'une fonction de deux variables

La continuité d'une fonction de deux variables se définit à l'aide de la notion de limite.

1. Définition de la continuité d'une fonction de deux variables. Soient $f(M)$ une fonction définie sur un ensemble $\{M\}$ et $M_0 \in \{M\}$ un point dont tout δ-voisinage contient des points de $\{M\}$ autres que M_0.

Définition 1. On dit que la fonction $z = f(M)$ est *continue au point* M_0 si elle admet une limite égale à sa valeur en ce point, c'est-à-dire que

$$\lim_{M \to M_0} f(M) = f(M_0) \quad \text{ou} \quad \lim_{\substack{x \to x_0 \\ y \to y_0}} f(x, y) = f(x_0, y_0).$$

D'après la définition de la limite d'une fonction en termes de suites, cette définition de la continuité d'une fonction en M_0 équivaut au fait que pour toute suite $\{M_n\}$ $(M_n \in \{M\})$ convergeant vers M_0, la suite $\{f(M_n)\}$ converge vers $f(M_0)$:

$$\lim_{n \to \infty} f(M_n) = f(M_0) \quad \text{ou} \quad \lim_{n \to \infty} f(x_n, y_n) = f(x_0, y_0).$$

Les points en lesquels une fonction n'est pas continue s'appellent *points de discontinuité* de cette fonction.

Par exemple, la fonction

$$f(x, y) = \begin{cases} \frac{x-y}{x+y} & \text{si } x+y \neq 0, \\ 1 & \text{si } x+y=0 \end{cases}$$

est discontinue en (0; 0), puisqu'elle n'admet pas de limite lorsque $x \to 0$, $y \to 0$; la fonction

$$f(x, y) = \begin{cases} x^2+y^2 & \text{partout sauf pour } x=1,\ y=2, \\ 0 & \text{pour } x=1,\ y=2 \end{cases}$$

est discontinue au point (1, 2), puisque $\lim\limits_{\substack{x \to 1 \\ y \to 2}} f(x, y) = 5$ et $f(1; 2) = 0$.

Définissons la continuité d'une fonction à l'aide de la définition de la limite d'une fonction en termes de ε et δ.

Définition 2. On dit qu'une fonction $z = f(M)$ est *continue en un point* M_0 si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout point $M \in \{M\}$ vérifiant la condition $\rho(M; M_0) < \delta$ l'on ait $|f(M) - f(M_0)| < \varepsilon$.

En signes logiques cette définition s'écrit

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall M \in \{M\},\ \rho(M; M_0) < \delta : |f(M) - f(M_0)| < \varepsilon.$$

En se servant de ces définitions et des théorèmes respectifs sur les limites on peut comme pour le cas d'une fonction d'une variable prouver que les opérations arithmétiques sur les fonctions continues et la composition de fonctions continues sont des opérations stables.

La définition 1 de la continuité d'une fonction est utilisée dans la suite sous une autre forme.

Appelons *accroissement total* de la fonction $z = f(M)$ en un point M_0 la fonction Δz définie par la formule

$$\Delta z = f(M) - f(M_0),$$

où M est un point quelconque du domaine de définition de z. Soient $(x_0; y_0)$ et $(x; y)$ les coordonnées respectives des points M_0 et M. Posons $x - x_0 = \Delta x$ et $y - y_0 = \Delta y$. Pour Δz on obtient alors l'expression suivante:

$$\Delta z = f(x_0 + \Delta x,\ y_0 + \Delta y) - f(x_0,\ y_0).$$

Définition 3. On dit que la fonction $z = f(M)$ est *continue en* M_0 si son accroissement total en M_0 est un infiniment petit lorsque $M \to M_0$, c'est-à-dire que

$$\lim_{M \to M_0} \Delta z = \lim_{M \to M_0} [f(M) - f(M_0)] = 0 \quad \text{ou} \quad \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \Delta z = 0.$$

Cette condition équivaut visiblement à la condition $\lim\limits_{M \to M_0} f(M) = f(M_0)$ de la définition 1.

Exemple. La fonction $z = x^2 + y^2$ est continue en tout point $(x; y)$. En effet, son accroissement total au point $(x; y)$ est

$$\Delta z = [(x + \Delta x)^2 + (y + \Delta y)^2] - (x^2 + y^2) = \\ = 2x\Delta x + 2y\Delta y + (\Delta x)^2 + (\Delta y)^2.$$

Il est évident que Δz tend vers 0 avec Δx et Δy, c'est-à-dire que la fonction $z = x^2 + y^2$ est continue au point $(x; y)$ aux termes de la définition 3.

On dit que la fonction $z = f(M)$ est continue sur un ensemble $\{M\}$ si elle l'est en tout point de cet ensemble.

2. Propriétés fondamentales des fonctions continues de deux variables. Citons sans les prouver les propriétés fondamentales

des fonctions continues de deux variables, puisque leurs démonstrations sont calquées sur celles des propriétés correspondantes des fonctions d'une variable. Introduisons préalablement quelques notions sur les ensembles $\{M\}$ de points du plan.

Définition 4. Un ensemble $\{M\}$ de points du plan est dit *connexe* s'il contient deux quelconques de ses points avec une ligne continue les reliant.

Exemple: un disque est un ensemble connexe; un ensemble composé de deux disques sans points communs n'est pas connexe.

Définition 5. On dit qu'un point M est un point *intérieur* d'un ensemble $\{M\}$ si tout δ-voisinage de M contient des points de $\{M\}$.

Définition 6. Un ensemble $\{M\}$ composé uniquement de points intérieurs s'appelle *ensemble ouvert.*

Définition 7. Un ensemble $\{M\}$ de points, ouvert et connexe, s'appelle *domaine ouvert* ou simplement *domaine.*

Des exemples élémentaires de domaines nous sont donnés par l'intérieur d'un triangle, d'un disque, d'une ellipse, etc.

Définition 8. Un point M s'appelle *point frontière* d'un domaine si tout δ-voisinage de M contient des points intérieurs et des points extérieurs à ce domaine. L'ensemble de tous les points frontières d'un domaine s'appelle *frontière* ou *bord* de ce domaine.

Par exemple, la frontière d'un domaine dont les points sont situés dans un disque est un cercle.

Définition 9. L'ensemble $\{M\}$ composé d'un domaine et de sa frontière s'appelle *domaine fermé.*

Définition 10. Un ensemble $\{M\}$ est dit *borné* s'il est entièrement contenu dans un disque.

Un segment et un triangle sont des ensembles bornés. La droite n'est pas un ensemble borné.

Les fonctions de deux variables se définissent sur un domaine fermé borné par analogie aux fonctions d'une variable dont le domaine de définition est un intervalle fermé borné, c'est-à-dire un segment.

Formulons maintenant les propriétés fondamentales des fonctions continues de deux variables.

1° Si une fonction $z = f(M)$ est continue sur un domaine fermé borné, elle est bornée sur ce domaine, c'est-à-dire qu'il existe un nombre k tel que $|f(M)| \leqslant k$ pour tout point de ce domaine.

2° Si une fonction $z = f(M)$ est continue sur un domaine fermé borné, elle atteint ses bornes supérieure et inférieure sur ce domaine.

3° Si une fonction $z = f(M)$ est continue sur un domaine, elle prend toutes les valeurs intermédiaires entre deux quelconques de ses valeurs, autrement dit si $A < C < B$, où A et B sont des valeurs de $f(M)$ sur ce domaine, il existe un point M_0 de ce domaine tel que $f(M_0) = C$.

Il s'ensuit en particulier que si M_1 et M_2 sont des points de ce domaine et si $f(M_1) < 0$ et $f(M_2) > 0$, alors il existe un point M_0 de ce domaine en lequel $f(M_0) = 0$.

4° Si une fonction $z = f(M)$ est continue sur un domaine fermé borné, elle est uniformément continue sur ce domaine, c'est-à-dire que pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout couple M' et M'' satisfaisant à la condition $\rho(M'; M'') < \delta$ l'on a $|f(M'') - f(M')| < \varepsilon$.

Signalons en conclusion que les notions de limite, de continuité et les propriétés citées des fonctions de deux variables se généralisent facilement aux fonctions de trois et de plus de trois variables.

CHAPITRE XII

DÉRIVÉES PARTIELLES ET DÉRIVABILITÉ DES FONCTIONS DE PLUSIEURS VARIABLES

§ 1. Dérivées partielles

Soit $z = f(M)$ une fonction définie sur un voisinage d'un point $M(x; y)$ *). Donnons à la variable x un accroissement arbitraire Δx en laissant y fixe, c'est-à-dire passons du point $M(x; y)$ au point $M_1(x + \Delta x; y)$. L'accroissement Δx est tel que le point M_1 reste dans le voisinage envisagé du point M. L'accroissement correspondant de la fonction

$$\Delta_x z = f(x + \Delta x, y) - f(x, y)$$

s'appelle *accroissement partiel de la fonction par rapport à la variable* x au point $M(x; y)$.

On définit de façon analogue l'*accroissement partiel de la fonction par rapport à* y:

$$\Delta_y z = f(x, y + \Delta y) - f(x, y).$$

Définition. La limite

$$\lim_{\Delta x \to 0} \frac{\Delta_x z}{\Delta x} \quad \left(\text{resp. } \lim_{\Delta y \to 0} \frac{\Delta_y z}{\Delta y}\right),$$

si elle existe s'appelle *dérivée partielle de la fonction* $z = f(M)$ *en* M *par rapport à la variable* x (resp. *à la variable* y) et se désigne par l'un des symboles suivants

$$z'_x,\ f'_x,\ \frac{\partial f}{\partial x},\ \frac{\partial z}{\partial x} \quad \left(\text{resp. } z'_y,\ f'_y,\ \frac{\partial f}{\partial y},\ \frac{\partial z}{\partial y}\right)$$

(on omet parfois l'accent et on écrit z_x).

Il résulte de la définition que la dérivée partielle d'une fonction de deux variables par rapport à x est une ordinaire dérivée de cette fonction traitée comme une fonction de la variable x à y fixe. Les dérivées partielles sont donc justiciables des mêmes formules et règles de calcul que les dérivées des fonctions d'une variable.

---

*) On appelle *voisinage* d'un point M un domaine contenant M. En particulier, si ce domaine est un disque ouvert de rayon δ et de centre M, ce voisinage est un δ-voisinage de M.

Exemples.

1. $z = x^2 - 2xy^2 + y^3$, $\frac{\partial z}{\partial x} = 2x - 2y^2$, $\frac{\partial z}{\partial y} = -4xy + 3y^2$.

2. $z = \operatorname{Arc\,tg} \frac{x}{y}$, $\frac{\partial z}{\partial x} = \frac{y}{x^2+y^2}$, $\frac{\partial z}{\partial y} = \frac{-x}{x^2+y^2}$.

3. $z = x^2 \sin y$, $\frac{\partial z}{\partial x} = 2x \sin y$, $\frac{\partial z}{\partial y} = x^2 \cos y$.

4. $z = xye^{x+2y}$, $\frac{\partial z}{\partial x} = e^{x+2y} y (1 + x)$, $\frac{\partial z}{\partial y} = e^{x+2y} x (1 + 2y)$.

A noter que nous avons défini les dérivées partielles de la fonction $z = f(x, y)$ en un point M au voisinage duquel cette fonction est définie, c'est-à-dire en un point intérieur du domaine de définition. Si $M(x; y)$ est un point frontière du domaine de définition de la fonction, alors $\Delta_x z$ (resp. $\Delta_y z$) peut ne pas être défini, puisque le point $M_1(x + \Delta x; y)$ (resp. $M_2(x; y + \Delta y)$) peut ne pas appartenir au domaine de définition pour aucun $\Delta x \neq 0$ (resp. $\Delta y \neq 0$). Ceci concerne par exemple le point M_0 de la figure 164. Dans ce cas si la dérivée partielle z'_x existe en tout point intérieur M du domaine et si existe $\lim_{M \to M_0} z'_x(M)$, on pose par définition

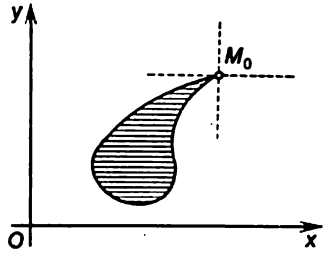


Fig. 164

$$z'_x(M_0) = \lim_{M \to M_0} z'_x(M).$$

On définit $z'_y(M_0)$ de façon analogue.

§ 2. Notion de dérivabilité d'une fonction

1. Définition de la dérivabilité. On rappelle que l'accroissement total d'une fonction $z = f(M)$ en un point $M(x; y)$, correspondant aux accroissements Δx et Δy des variables x et y est la fonction

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y).$$

Soit $z = f(M)$ une fonction définie sur un voisinage d'un point M.

Définition. On dit que la fonction $z = f(M)$ est *dérivable en* M si son accroissement total en M peut être représenté sous la forme

$$\Delta z = A \Delta x + B \Delta y + \alpha(\Delta x, \Delta y) \Delta x + \beta(\Delta x, \Delta y) \Delta y, \quad (1)$$

où A et B sont des nombres indépendants de Δx et Δy et $\alpha(\Delta x, \Delta y)$ et $\beta(\Delta x, \Delta y)$ des infiniment petits lorsque $\Delta x \to 0$, $\Delta y \to 0$.

On sait que si une fonction d'une variable est dérivable en un point, elle est continue et possède une dérivée en ce point. L'existence de la dérivée d'une fonction d'une variable en un point entraîne la dérivabilité de cette fonction en ce point. Voyons comment ces propriétés se généralisent à une fonction de deux variables.

2. Conditions nécessaires de dérivabilité.

Théorème 12.1. *Si une fonction $z = f(M)$ est dérivable en un point M, elle est continue en ce point.*

Démonstration. Si $z = f(M)$ est dérivable en M, la relation (1) entraîne $\lim\limits_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \Delta z = 0$, ce qui exprime que cette fonction est continue en M. ■

Théorème 12.2. *Si une fonction $z = f(M)$ est dérivable en $M(x; y)$, elle admet en ce point les dérivées partielles $f'_x(x, y)$ et $f'_y(x, y)$ et de plus*

$$f'_x(x, y) = A, \quad f'_y(x, y) = B.$$

Démonstration. La fonction $z = f(M)$ étant dérivable en M, on a la relation (1). En faisant $\Delta y = 0$, on obtient $\Delta_x z = A\Delta x + \alpha(\Delta x, 0)\Delta x$, où $\alpha(\Delta x, 0)$ est un infiniment petit lorsque $\Delta x \to 0$. En divisant par Δx et en faisant tendre Δx vers 0 on obtient

$$\lim_{\Delta x \to 0} \frac{\Delta_x z}{\Delta x} = \lim_{\Delta x \to 0} [A + \alpha(\Delta x, 0)] = A.$$

Donc la dérivée partielle par rapport à x existe en M et $f'_x(x, y) = A$.

On démontre de façon analogue que la dérivée partielle par rapport à y existe en M et $f'_y(x, y) = B$. ■

Les réciproques des théorèmes 12.1 et 12.2 sont fausses, autrement dit, ni la continuité d'une fonction de deux variables en un point M, ni l'existence de ses dérivées partielles en ce point n'entraînent sa dérivabilité.

Par exemple, la fonction $f(x, y) = \sqrt{x^2 + y^2}$ est continue en $(0; 0)$, mais elle ne possède pas de dérivées partielles en ce point. En effet,

$$\frac{f(0+\Delta x, 0) - f(0, 0)}{\Delta x} = \frac{\sqrt{(0+\Delta x)^2 + 0} - 0}{\Delta x} = \frac{|\Delta x|}{\Delta x}.$$

Or la fonction $\frac{|\Delta x|}{\Delta x}$ n'admet pas de limite lorsque $\Delta x \to 0$. Donc $f'_x(0, 0)$ n'existe pas. On démontre de façon analogue que

$f'_y(0, 0)$ n'existe pas non plus. Cette fonction n'est pas dérivable en ce point, puisqu'elle n'admet pas de dérivées partielles.

La fonction

$$f(x, y) = \begin{cases} 0 & \text{sur les axes de coordonnées,} \\ 1 & \text{ailleurs} \end{cases}$$

possède des dérivées partielles par rapport à x et y au point (0 ; 0). Ceci résulte du fait que $f(x, 0) \equiv 0$ et $f(0, y) \equiv 0$, donc $f'_x(0, 0) = 0$ et $f'_y(0, 0) = 0$. Mais $f(x, y)$ n'est pas continue en ce point, puisque $\lim\limits_{\substack{x\to 0 \\ y\to 0}} f(x, y) = 1$ sur la droite $y = x$ et $f(0, 0) = 0$. Donc $f(x, y)$ n'est pas dérivable en (0 ; 0).

Considérons un dernier exemple : la fonction $f(x, y) = \sqrt{|x|\,|y|}$ est continue en (0; 0), puisque $f(0, 0) = 0$, $\lim\limits_{\substack{x\to 0 \\ y\to 0}} f(x, y) = 0$ et

$$f'_x(0, 0) = \lim_{\Delta x\to 0} \frac{\sqrt{|\Delta x|\cdot 0} - 0}{\Delta x} = 0,$$

$$f'_y(0, 0) = \lim_{\Delta y\to 0} \frac{\sqrt{0\cdot|\Delta y|} - 0}{\Delta y} = 0,$$

pourtant cette fonction n'est pas dérivable au point (0; 0). En effet, son accroissement total en (0; 0) est

$$\Delta z = \sqrt{|\Delta x|\,|\Delta y|}.$$

Si cette fonction était dérivable en (0; 0), on aurait en vertu de la relation (1) et du théorème 12.2

$$\Delta z = 0\cdot\Delta x + 0\cdot\Delta y + \alpha\Delta x + \beta\Delta y,$$

où α et $\beta \to 0$ avec Δx et Δy. D'autre part, de l'égalité

$$\alpha\,\Delta x + \beta\,\Delta y = \left(\frac{\alpha\,\Delta x + \beta\,\Delta y}{\rho}\right)\rho = \gamma\rho,$$

où

$$\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2},$$

on déduit que γ tend vers 0 avec ρ [cf. § 4, formule (3)]. Mais dans ce cas, pour tout $\Delta x = \Delta y$

$$\gamma = \sqrt{|\Delta x|\,|\Delta y|}\,/\,\sqrt{(\Delta x)^2 + (\Delta y)^2} = 1/\sqrt{2},$$

c'est-à-dire que γ n'est pas un infiniment petit lorsque $\rho \to 0$.

Donc en (0 ; 0) la fonction $f(x, y) = \sqrt{|x|\,|y|}$ est continue, possède des dérivées partielles et, pourtant, elle n'est pas dérivable.

3. Conditions suffisantes de dérivabilité.

Théorème 12.3. *Si une fonction $z = f(M)$ possède des dérivées partielles sur un δ-voisinage d'un point M et si ces dérivées sont continues en M, cette fonction est dérivable en M.*

Démonstration. Donnons aux variables x et y des accroissements Δx et Δy assez petits pour que le point $M_1(x+\Delta x; y+\Delta y)$ reste dans ce δ-voisinage de M. L'accroissement total de z

$$\Delta z = f(x+\Delta x, y+\Delta y) - f(x, y)$$

peut être mis sous la forme

$$\Delta z = [f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)] + [f(x, y+\Delta y) - f(x, y)]. \quad (2)$$

L'expression $[f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)]$ peut être traitée comme l'accroissement de la fonction $f(x, y+\Delta y)$ de la variable x à $y+\Delta y$ constant. Cette fonction possédant la dérivée $f'_x(x, y+\Delta y)$ par hypothèse, le théorème de Lagrange nous donne

$$f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) = f'_x(x+\theta_1\Delta x, y+\Delta y)\,\Delta x, \quad 0<\theta_1<1.$$

Mutatis mutandis, on obtient

$$f(x, y+\Delta y) - f(x, y) = f'_y(x, y+\theta_2\Delta y)\,\Delta y, \quad 0<\theta_2<1.$$

Les dérivées f'_x et f'_y étant continues en $M(x; y)$, il vient

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} f'_x(x+\theta_1\Delta x, y+\Delta y) = f'_x(x, y),$$

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} f'_y(x, y+\theta_2\Delta y) = f'_y(x, y).$$

D'où

$$f'_x(x+\theta_1\Delta x, y+\Delta y) = f'_x(x, y) + \alpha(\Delta x, \Delta y),$$

$$f'_y(x, y+\theta_2\Delta y) = f'_y(x, y) + \beta(\Delta x, \Delta y),$$

où $\alpha(\Delta x, \Delta y)$ et $\beta(\Delta x, \Delta y)$ sont des infiniment petits lorsque $\Delta x\to 0$ et $\Delta y\to 0$. En portant les expressions obtenues dans la formule (2) de Δz, on trouve

$$\Delta z = f'_x(x, y)\,\Delta x + f'_y(x, y)\,\Delta y + \alpha(\Delta x, \Delta y)\,\Delta x + \beta(\Delta x, \Delta y)\,\Delta y,$$

ce qui exprime que la fonction $z = f(M)$ est dérivable en M. ■

Corollaire. *La continuité des dérivées partielles entraîne celle de la fonction.*

Le théorème 12.3 est d'une grande importance pour l'établissement de la dérivabilité d'une fonction, car la vérification

directe de la dérivabilité d'une fonction à l'aide de la définition est souvent très compliquée, tandis qu'il est plus facile de dire si les dérivées partielles de cette fonction sont continues.

Signalons en conclusion que la dérivabilité d'une fonction de trois et de plus de trois variables se définit comme pour une fonction de deux variables.

§ 3. Dérivées de fonctions composées

Soit $z = f(x, y)$ une fonction de deux variables x et y qui à leur tour sont fonctions d'une variable indépendante t: $x = x(t)$, $y = y(t)$. La fonction $z = f[x(t), y(t)]$ est alors une fonction composée de la variable indépendante t, les variables x et y sont des variables intermédiaires. On a le théorème suivant.

Théorème 12.4. *Si des fonctions $x = x(t)$ et $y = y(t)$ sont dérivables en un point t et si une fonction $z = f(x, y)$ est dérivable en un point $M[x(t); y(t)]$, alors la fonction composée $z = f[(x(t), y(t)]$ est dérivable aussi au point t et de plus*

$$\frac{dz}{dt} = \frac{dz}{dx}\frac{dx}{dt} + \frac{dz}{dy}\frac{dy}{dt}. \tag{1}$$

Démonstration. Donnons à la variable t un accroissement arbitraire Δt; les fonctions $x(t)$ et $y(t)$ subissent les accroissements Δx et Δy et la fonction $z = f(x, y)$, l'accroissement

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y).$$

La fonction $z = f(x, y)$ étant dérivable au point $M(x; y)$, où $x = x(t)$, $y = y(t)$, l'accroissement Δz peut être mis sous la forme

$$\Delta z = f'_x(x, y)\,\Delta x + f'_y(x, y)\,\Delta y + \alpha(\Delta x, \Delta y)\,\Delta x + \beta(\Delta x, \Delta y)\,\Delta y,$$

où $\alpha(\Delta x, \Delta y)$ et $\beta(\Delta x, \Delta y)$ sont des infiniment petits lorsque $\Delta x \to 0$, $\Delta y = 0$. Prolongeons ces fonctions en posant $\alpha(0, 0) = 0$, $\beta(0, 0) = 0$.

En divisant les deux membres de l'expression de Δz par Δt, on obtient

$$\frac{\Delta z}{\Delta t} = f'_x(x, y)\frac{\Delta x}{\Delta t} + f'_y(x, y)\frac{\Delta y}{\Delta t} + \alpha(\Delta x, \Delta y)\frac{\Delta x}{\Delta t} + \beta(\Delta x, \Delta y)\frac{\Delta y}{\Delta t}. \tag{2}$$

Par hypothèse, $\lim\limits_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = \frac{dx}{dt}$, $\lim\limits_{\Delta t \to 0} \frac{\Delta y}{\Delta t} = \frac{dy}{dt}$. Par ailleurs, étant dérivables au point t, les fonctions $x(t)$ et $y(t)$ sont continues en ce point, c'est-à-dire que Δx et Δy tendent vers 0

avec Δt et par suite $\alpha(\Delta x, \Delta y) \to 0$ et $\beta(\Delta x, \Delta y) \to 0$. Donc $\alpha(\Delta x, \Delta y)\frac{\Delta x}{\Delta t}$ et $\beta(\Delta x, \Delta y)\frac{\Delta y}{\Delta t}$ tendent vers 0 avec Δt.

Nous avons ainsi montré que le second membre de (2) possède une limite lorsque $\Delta t \to 0$, donc il en est de même du premier membre :

$$\lim_{\Delta t \to 0} \frac{\Delta z}{\Delta t} = \frac{dz}{dt},$$

et de plus

$$\frac{dz}{dt} = f'_x(x, y)\frac{dx}{dt} + f'_y(x, y)\frac{dy}{dt} \quad \text{ou} \quad \frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}. \blacksquare$$

Exemples.

1. Soient $z = f(x, y)$, $x = t^3 + 2$, $y = 3t^4 - 1$. La formule (1) nous donne

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} 3t^2 + \frac{\partial z}{\partial y} 12t^3.$$

2. Soient $z = x \sin\frac{x}{y}$, $x = 1 + 3t$, $y = \sqrt{1+t^2}$. La formule (1) nous donne

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} = \left(\sin\frac{x}{y} + \frac{x}{y}\cos\frac{x}{y}\right)3 - \frac{x^2}{y^2}\cos\frac{x}{y}\frac{t}{\sqrt{1+t^2}}.$$

3. Soient $z = x^2y^3$, $x = t$, $y = t^2$. D'après la formule (1)

$$\frac{dz}{dt} = 2xy^3 \cdot 1 + 3x^2y^2 \cdot 2t.$$

Comme $x = t$, $y = t^2$, on a $\frac{dz}{dt} = 2t(t^2)^3 + 3t^2(t^2)^2 2t = 8t^7$. D'autre part, on peut trouver $\frac{dz}{dt}$ en exprimant préalablement z en fonction de t. On a $z = x^2y^3 = t^2(t^2)^3 = t^8$, d'où $\frac{dz}{dt} = 8t^7$. On retrouve le résultat acquis par la formule (1).

Si $z = f(x, y)$, où $y = \varphi(x)$, alors $z = f[x, \varphi(x)]$ est une fonction composée de x. D'après la formule (1) où il faut remplacer t par x, on obtient

$$\frac{dz}{dx} = \frac{\partial z}{\partial x}\frac{dx}{dx} + \frac{\partial z}{\partial y}\frac{dy}{dx},$$

et comme $\frac{dx}{dx} = 1$,

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx}.$$

On envisage de façon analogue le problème de la dérivée d'une fonction composée lorsque les variables intermédiaires sont plus de deux. Par exemple, si $u = f(x, y, z)$, où $x = x(t)$, $y = y(t)$, $z = z(t)$, la formule (1) devient

$$\frac{du}{dt} = \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt} + \frac{\partial u}{\partial z}\frac{dz}{dt}.$$

Considérons maintenant un cas plus général. Soit $z = f(x, y)$ une fonction de deux variables x et y, dépendant à leur tour de deux ou d'un plus grand nombre de variables indépendantes. Supposons, par exemple, que $x = x(u, v)$, $y = y(u, v)$. La fonction $z = f[x(u, v), y(u, v)]$ est alors une fonction composée des variables indépendantes u et v, les variables x et y étant intermédiaires.

Si les fonctions $x(u, v)$ et $y(u, v)$ sont dérivables en un point $M'(u; v)$ et la fonction $z = f(x, y)$, dérivable en $M(x; y)$, où $x = x(u, v)$, $y = y(u, v)$, la fonction composée $z = f[(x(u, v), y(u, v)]$ est dérivable en $M'(u, v)$ et ses dérivées partielles en ce point sont données par les formules

$$\begin{cases} \dfrac{\partial z}{\partial u} = \dfrac{\partial z}{\partial x}\dfrac{\partial x}{\partial u} + \dfrac{\partial z}{\partial y}\dfrac{\partial y}{\partial u}, \\ \dfrac{\partial z}{\partial v} = \dfrac{\partial z}{\partial x}\dfrac{\partial x}{\partial v} + \dfrac{\partial z}{\partial y}\dfrac{\partial y}{\partial v}. \end{cases} \tag{3}$$

Nous glisserons sur la démonstration de la dérivabilité d'une fonction composée. Pour calculer ses dérivées partielles, on fixe la valeur de l'une des variables u ou v. On se place alors dans les conditions du théorème 12.4 qui vient tout juste d'être prouvé. Les formules (3) résultent de la formule (1) dans les deux cas.

Exemples.

1. Soient $z = f(x, y)$, $x = u^2 + 2v$, $y = \dfrac{u^2}{v}$. Les formules (3) nous donnent

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}2u + \frac{\partial z}{\partial y}\frac{2u}{v}; \quad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}2 + \frac{\partial z}{\partial y}\left(-\frac{u^2}{v^2}\right).$$

2. Soient $z = x^2y^2$, $x = u + v$, $y = \dfrac{u}{v}$. D'après les formules (3)

$$\frac{\partial z}{\partial u} = 2xy^2 \cdot 1 + 2x^2y \cdot \frac{1}{v}; \quad \frac{\partial z}{\partial v} = 2xy^2 \cdot 1 + 2x^2y \cdot \left(-\frac{u}{v^2}\right).$$

Porter les expressions $x = u + v$, $y = \dfrac{u}{v}$ dans ces formules ; trouver $\dfrac{\partial z}{\partial u}$ et $\dfrac{\partial z}{\partial v}$ après avoir exprimé préalablement z en fonction de u et v ; comparer ensuite les résultats obtenus.

3. Soient $z = x^2 - y^2$, $x = u \cos v$, $y = u \sin v$. Les formules (3) nous donnent

$$\frac{\partial z}{\partial u} = 2x \cos v - 2y \sin v\,; \qquad \frac{\partial z}{\partial v} = -2xu \sin v - 2yu \cos v.$$

Si $z = f(x)$, où $x = x(u, v)$, alors $z = f[x(u, v)]$ est une fonction composée dépendant des variables u, v par l'intermédiaire de x, et ses dérivées partielles sont données aussi par les formules (3)

$$\begin{cases} \dfrac{\partial z}{\partial u} = \dfrac{dz}{dx}\dfrac{\partial x}{\partial u}, \\ \dfrac{\partial z}{\partial v} = \dfrac{dz}{dx}\dfrac{\partial x}{\partial v}. \end{cases}$$

Les formules (3) peuvent être généralisées à un plus grand nombre de variables intermédiaires.

Si, par exemple, $w = f(x, y, z)$ est une fonction de trois variables x, y, z dépendant chacune de u et v, les formules (3) deviennent

$$\begin{cases} \dfrac{\partial w}{\partial u} = \dfrac{\partial w}{\partial x}\dfrac{\partial x}{\partial u} + \dfrac{\partial w}{\partial y}\dfrac{\partial y}{\partial u} + \dfrac{\partial w}{\partial z}\dfrac{\partial z}{\partial u}, \\ \dfrac{\partial w}{\partial v} = \dfrac{\partial w}{\partial x}\dfrac{\partial x}{\partial v} + \dfrac{\partial w}{\partial y}\dfrac{\partial y}{\partial v} + \dfrac{\partial w}{\partial z}\dfrac{\partial z}{\partial v}. \end{cases}$$

§ 4. Différentielle d'une fonction

1. Définition de la différentielle. On rappelle: si une fonction $z = f(M)$ est dérivable en un point M, son accroissement total en ce point peut être représenté sous la forme

$$\Delta z = A\Delta x + B\Delta y + \alpha(\Delta x, \Delta y)\,\Delta x + \beta(\Delta x, \Delta y)\,\Delta y, \qquad (1)$$

où $\alpha(\Delta x, \Delta y)$ et $\beta(\Delta x, \Delta y)$ sont des infiniment petits lorsque $\Delta x \to 0$, $\Delta y \to 0$.

Définition. On appelle *différentielle* dz d'une fonction $z = f(M)$ dérivable en un point M la partie linéaire en Δx et Δy de l'accroissement de cette fonction en M, c'est-à-dire

$$dz = A\Delta x + B\Delta y. \qquad (2)$$

Le théorème 12.2 nous permet de mettre l'expression (2) sous la forme

$$dz = f'_x(x, y)\,\Delta x + f'_y(x, y)\,\Delta y.$$

Appelons *différentielles des variables indépendantes* x et y les accroissements de ces variables: $dx = \Delta x$, $dy = \Delta y$. La différentielle de la fonction s'écrit alors

$$dz = f'_x(x, y)\,dx + f'_y(x, y)\,dy.$$

Des relations (1) et (2) il s'ensuit que la différence entre l'accroissement total et la différentielle de z en M, soit

$$\Delta z - dz = \alpha(\Delta x, \Delta y)\,\Delta x + \beta(\Delta x, \Delta y)\,\Delta y$$

est un infiniment petit d'ordre supérieur à $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$ [ρ est la distance des points $M(x; y)$ et $M_1(x + \Delta x; y + \Delta y)$]. En effet,

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{\Delta z - dz}{\rho} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \left(\alpha \frac{\Delta x}{\rho} + \beta \frac{\Delta y}{\rho}\right) = 0,$$

puisque α et β sont des infiniment petits et $\frac{\Delta x}{\rho}$, $\frac{\Delta y}{\rho}$, des fonctions bornées : $\left|\frac{\Delta x}{\rho}\right| \leqslant 1$, $\left|\frac{\Delta y}{\rho}\right| \leqslant 1$. On en déduit que $\Delta z - dz = o(\rho)$ ou

$$\Delta z = dz + o\left(\sqrt{(\Delta x)^2 + (\Delta y)^2}\right).$$

En négligeant la quantité $o(\rho)$ pour Δx et Δy assez petits, on obtient la formule approchée

$$\Delta z \approx dz,$$

qui est largement utilisée dans les calculs approchés, puisqu'il est plus facile de calculer une différentielle qu'un accroissement total.

2. Plan tangent et normale à une surface. Interprétation géométrique de la différentielle. De même que la différentielle d'une fonction d'une variable se représente géométriquement par l'accroissement de l'« ordonnée d'une tangente », de même la différentielle d'une fonction de deux variables est l'accroissement de la « cote d'un plan tangent ». Introduisons la notion de plan tangent à une surface en un point N_0.

On dit qu'un plan passant par un point N_0 d'une surface est un *plan tangent en* N_0 à cette surface si l'angle qu'il fait avec une sécante passant par N_0 et par un point quelconque N de la surface tend vers 0 lorsque N tend vers N_0.

Supposons qu'une surface est définie par l'équation $z = f(x, y)$, et que la fonction $f(x, y)$ est dérivable au point $M_0(x_0; y_0)$.

Montrons que le plan tangent à la surface au point $N_0(x_0; y_0; z_0)$, où $z_0 = f(x; y)$ est représenté par l'équation

$$z - z_0 = f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0). \qquad (4)$$

En effet, du cours de géométrie analytique on sait que l'équation (4) représente le plan passant par le point $N_0(x_0; y_0; z_0)$, de vecteur normal $\vec{n} = \{f'_x; f'_y; -1\}$. Pour montrer que ce plan est tangent, il suffit de prouver que l'angle φ du vecteur $\vec{n}$ et du vec-

теур директеур $\overrightarrow{N_0N}$ де ла сécante N_0N тенд верс $\pi/2$ лорсque N тенд верс N_0. Дéсигнонс лес коордоннéес де N пар $(x; y; z)$, où $z = f(x, y)$. Пуисque $\vec{n} = \{f'_x; f'_y; -1\}$ ет $\overrightarrow{N_0N} = \{x - x_0; y - y_0; z - z_0\}$, он а

$$\cos\varphi = \frac{f'_x(x-x_0)+f'_y(y-y_0)-(z-z_0)}{\sqrt{f'^2_x+f'^2_y+1}\,\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}}.$$

Ор де ла релатион (3) иль с'енсуит que $f'_x(x-x_0)+f'_y(y-y_0)-(z-z_0)=o(\rho)$, où $\rho=\sqrt{(x-x_0)^2+(y-y_0)^2}$. Донс

$$|\cos\varphi| \leqslant \frac{|o(\rho)|}{\sqrt{(x-x_0)^2+(y-y_0)^2}} = \frac{|o(\rho)|}{\rho} \to 0,$$

лорсque $\rho \to 0$. D'où $\lim\limits_{N\to N_0} \varphi = \pi/2$. C.Q.F.D.

Le vecteur normal $\vec{n} = \{f'_x; f'_y; -1\}$ du plan tangent s'appelle *normale* à la surface $z = f(x, y)$ en N_0. Soient $x - x_0 = \Delta x$, $y - y_0 = \Delta y$, $z - z_0 = \Delta z$. De l'équation (4) on déduit alors que l'accroissement Δz de la «cote du plan tangent» est défini par la formule

$$\Delta z = f'_x(x_0, y_0)\,\Delta x + f'_y(x_0, y_0)\,\Delta y,$$

qui est confondue avec l'expression de la différentielle dz de la fonction $z = f(x, y)$.

§ 5. Dérivée suivant une direction. Gradient

Soient $z = f(M)$ une fonction définie au voisinage d'un point $M(x; y)$ et un vecteur unitaire arbitraire $\vec{l} = \{\cos\alpha; \cos\beta\}$ (fig. 165).

Introduisons la notion de dérivée suivant une direction pour caractériser la vitesse de variation de la fonction z en $M(x; y)$ suivant la direction du vecteur $\vec{l}$. A cet effet, menons par le point M une droite L de même direction et de même sens que $\vec{l}$ et soit $M_1(x + \Delta x; y + \Delta y)$ un point de cette droite. Désignons par Δl la mesure algébrique du segment $\overrightarrow{MM_1}$. Dans le cas de figure $\Delta l = \sqrt{(\Delta x)^2 + (\Delta y)^2}$. L'accroissement correspondant de la fonc-

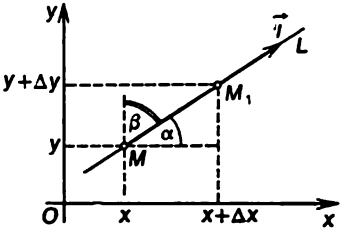


Fig. 165

tion $f(M)$ est

$$\Delta z = f(x + \Delta x,\ y + \Delta y) - f(x,\ y).$$

Définition 1. La limite, si elle existe, du rappot $\frac{\Delta z}{\Delta l}$ lorsque $\Delta l \to 0$ ($M_1 \to M$) s'appelle *dérivée de la fonction* $z = f(M)$ *au point* $M(x\,;\,y)$ *suivant la direction du vecteur* $\vec{l}$ et se note $\frac{\partial z}{\partial l}$, c'est-à-dire que

$$\lim_{\Delta l \to 0} \frac{\Delta z}{\Delta l} = \frac{\partial z}{\partial l}.$$

Supposons maintenant que la fonction $f(M)$ est dérivable en M. Son accroissement en ce point suivant la droite L peut être mis sous la forme

$$\Delta z = f'_x(x,\ y)\,\Delta x + f'_y(x,\ y)\,\Delta y + \alpha_1(\Delta x,\ \Delta y)\,\Delta x + \beta_1(\Delta x,\ \Delta y)\,\Delta y,$$

où α_1 et β_1 sont des infiniment petits lorsque $\Delta l \to 0$. En divisant les deux membres de cette égalité par Δl et puisque $\Delta x = \Delta l \cos \alpha$, $\Delta y = \Delta l \sin \alpha = \Delta l \cos \beta$, on obtient

$$\frac{\Delta z}{\Delta l} = f'_x(x,\ y)\cos\alpha + f'_y(x,\ y)\cos\beta + \\ + \alpha_1(\Delta x,\ \Delta y)\cos\alpha + \beta_1(\Delta x,\ \Delta y)\cos\beta.$$

En faisant tendre Δl vers 0, on obtient la formule de la dérivée suivant $\vec{l}$:

$$\frac{\partial z}{\partial l} = \frac{\partial z}{\partial x}\cos\alpha + \frac{\partial z}{\partial y}\cos\beta. \qquad (1)$$

On voit sur la formule (1) que la dérivée suivant $\vec{l}$ est une combinaison linéaire des dérivées partielles dans laquelle les cosinus directeurs semblent jouer le rôle de poids indiquant la contribution de la dérivée partielle correspondante.

En particulier, $\frac{\partial z}{\partial l} = \frac{\partial z}{\partial x}$ pour $\alpha = 0$ et $\beta = \frac{\pi}{2}$, $\frac{\partial z}{\partial l} = \frac{\partial z}{\partial y}$ pour $\alpha = \frac{\pi}{2}$ et $\beta = 0$. D'où il résulte que les dérivées partielles par rapport à x et à y sont des cas particuliers de la dérivée suivant une direction.

Exemple. Calculer la dérivée de la fonction $z = x^2 + y^2x$ au point M (1 ; 2) suivant le vecteur $\overrightarrow{MM_1}$, où M_1 est le point de coordonnées (3 ; 0).

Solution. Trouvons un vecteur unitaire $\vec{l}$ porté par $\overrightarrow{MM_1}$:

$$\overrightarrow{MM_1} = \{2;\ -2\} = 2\vec{i} - 2\vec{j};\quad |\overrightarrow{MM_1}| = 2\sqrt{2};$$

$$\vec{l} = \frac{\overrightarrow{MM_1}}{|\overrightarrow{MM_1}|} = \frac{2\vec{i} - 2\vec{j}}{2\sqrt{2}} = \frac{\sqrt{2}}{2}\vec{i} - \frac{\sqrt{2}}{2}\vec{j},$$

d'où $\cos\alpha = \sqrt{2}/2$, $\cos\beta = -\sqrt{2}/2$. Calculons les dérivées partielles de la fonction en M (1; 2). On a

$$f'_x(x, y) = 2x + y^2,\quad f'_y(x, y) = 2xy,$$

d'où

$$f'_x(1, 2) = 6,\quad f'_y(1, 2) = 4.$$

D'après la formule (1)

$$\frac{\partial z}{\partial l} = 6 \cdot \frac{\sqrt{2}}{2} - 4 \cdot \frac{\sqrt{2}}{2} = \sqrt{2}.$$

Définition 2. On appelle *gradient d'une fonction* $z = f(M)$ *en* $M(x; y)$ le vecteur dont les composantes sont égales respectivement aux dérivées partielles $\frac{\partial z}{\partial x}$ et $\frac{\partial z}{\partial y}$ prises en $M(x; y)$.

On note : $\overrightarrow{\text{grad}\, z} = \left\{\frac{\partial z}{\partial x};\ \frac{\partial z}{\partial y}\right\}$.

La notion de gradient nous permet de représenter la formule (1) par le produit scalaire des vecteurs $\overrightarrow{\text{grad}\, z}$ et $\vec{l}$:

$$\frac{\partial z}{\partial l} = \frac{\partial z}{\partial x}\cos\alpha + \frac{\partial z}{\partial y}\cos\beta = \overrightarrow{\text{grad}\, z} \cdot \vec{l}. \tag{2}$$

D'autre part, par définition du produit scalaire,

$$\overrightarrow{\text{grad}\, z} \cdot \vec{l} = |\overrightarrow{\text{grad}\, z}| \cdot |\vec{l}| \cos\varphi, \tag{3}$$

où $|\overrightarrow{\text{grad}\, z}|$ est le module du vecteur $\overrightarrow{\text{grad}\, z}$; φ, l'angle des vecteurs $\vec{l}$ et $\overrightarrow{\text{grad}\, z}$. En comparant les formules (2) et (3) et tenant compte du fait que $|\vec{l}| = 1$, on obtient

$$\frac{\partial z}{\partial l} = |\overrightarrow{\text{grad}\, z}| \cos\varphi.$$

De cette dernière égalité il s'ensuit que la dérivée suivant $\vec{l}$ prend sa plus grande valeur lorsque $\cos\varphi = 1$ ($\varphi = 0$), c'est-à-dire lorsque les vecteurs $\vec{l}$ et $\overrightarrow{\text{grad}\, z}$ sont de même direction. Dans ce cas $\frac{\partial z}{\partial l} = \overrightarrow{\text{grad}\, z}$.

Donc le gradient de la fonction $z = f(M)$ en $M(x; y)$ caractérise la direction et la valeur de la vitesse maximale de croissance de cette fonction en $M(x; y)$.

De façon analogue, on définit la dérivée suivant une direction d'une fonction $u = f(x, y, z)$ à trois variables. On établit la formule

$$\frac{\partial u}{\partial l} = \frac{\partial u}{\partial x}\cos\alpha + \frac{\partial u}{\partial y}\cos\beta + \frac{\partial u}{\partial z}\cos\gamma,$$

on introduit la notion de gradient

$$\overrightarrow{\operatorname{grad} u} = \left\{\frac{\partial u}{\partial x}; \frac{\partial u}{\partial y}; \frac{\partial u}{\partial z}\right\}$$

et on étudie ses propriétés.

Les notions de dérivée suivant une direction et de gradient d'une fonction sont très importantes dans les applications.

§ 6. Dérivées partielles et différentielles supérieures

1. Dérivées partielles supérieures. Supposons que les dérivées partielles $f'_x(x, y)$ et $f'_y(x, y)$ d'une fonction $z = f(M)$, définie au voisinage d'un point M, existent en tout point de ce voisinage. Dans ce cas ces dérivées sont des fonctions de x, y définies sur ce voisinage de M. Appelons-les *dérivées partielles premières* ou *du premier ordre*.

Les dérivées partielles par rapport à x et à y des fonctions $f'_x(x, y)$ et $f'_y(x, y)$ en M, si elles existent, s'appellent *dérivées partielles secondes* ou *du deuxième ordre* de la fonction $f(M)$ en M et se notent:

$$\frac{\partial^2 z}{\partial x^2} = f''_{xx}(x, y) = f^{(2)}_{x^2}(x, y); \quad \frac{\partial^2 z}{\partial y\,\partial x} = f''_{yx}(x, y) = f^{(2)}_{yx}(x, y);$$

$$\frac{\partial^2 z}{\partial y^2} = f''_{yy}(x, y) = f^{(2)}_{y^2}(x, y); \quad \frac{\partial^2 z}{\partial x\,\partial y} = f''_{xy}(x, y) = f^{(2)}_{xy}(x, y).$$

Les dérivées partielles $f''_{yx}(x, y)$ et $f''_{xy}(x, y)$ s'appellent *dérivées partielles mixtes*.

Exemples:

1. $z = x^4 + 4x^2y^3 + 7xy + 1$. On a

$$\frac{\partial z}{\partial x} = 4x^3 + 8xy^3 + 7y, \quad \frac{\partial z}{\partial y} = 12x^2y^2 + 7x.$$

Donc

$$\frac{\partial^2 z}{\partial x^2} = 12x^2 + 8y^3, \quad \frac{\partial^2 z}{\partial x\,\partial y} = \frac{\partial^2 z}{\partial y\,\partial x} = 24xy^2 + 7, \quad \frac{\partial^2 z}{\partial y^2} = 24x^2y.$$

2. $z = \sin x \cos y$. On a

$$\frac{\partial z}{\partial x} = \cos x \cos y, \qquad \frac{\partial z}{\partial y} = -\sin x \sin y.$$

Donc

$$\frac{\partial^2 z}{\partial x^2} = -\sin x \cos y, \qquad \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x} = -\cos x \sin y,$$

$$\frac{\partial^2 z}{\partial y^2} = -\sin x \cos y.$$

Les dérivées partielles mixtes $f''_{yx}(x, y)$ et $f''_{xy}(x, y)$ sont égales dans les deux exemples. Mais en général les valeurs des dérivées partielles mixtes dépendent de l'ordre de dérivation. Par exemple, la fonction

$$f(x, y) = \begin{cases} xy\,\dfrac{x^2 - y^2}{x^2 + y^2} & \text{si} \quad x^2 + y^2 \neq 0, \\ 0 & \text{si} \quad x^2 + y^2 = 0 \end{cases}$$

admet des dérivées partielles mixtes $f''_{xy}(x, y)$ et $f''_{yx}(x, y)$ en (0 ; 0) qui ne sont pas égales. En effet,

$$f'_x(x, y) = \begin{cases} \dfrac{y(x^4 - y^4 + 4x^2y^2)}{(x^2 + y^2)^2} & \text{si} \quad x^2 + y^2 \neq 0, \\ 0 & \text{si} \quad x^2 + y^2 = 0. \end{cases}$$

Donc

$$f''_{yx}(0, 0) = \lim_{\Delta y \to 0} \frac{f'_x(0, 0 + \Delta y) - f'_x(0, 0)}{\Delta y} = -1.$$

Des calculs analogues nous donnent: $f''_{xy}(0, 0) = 1$. Donc $f''_{yx}(0, 0) \neq f''_{xy}(0, 0)$.

A la question de savoir sous quelles conditions les valeurs des dérivées mixtes ne dépendent pas de l'ordre de dérivation répond le théorème suivant.

Théorème 12.5. *Si les dérivées $f''_{xy}(x, y)$ et $f''_{yx}(x, y)$ existent dans un δ-voisinage d'un point $M(x; y)$ et sont continues en M, elles sont égales en M, c'est-à-dire que*

$$f''_{xy}(x, y) = f''_{yx}(x, y).$$

Démonstration. Considérons l'expression

$$A = [f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y)] - [f(x, y + \Delta y) - f(x, y)],$$

où Δx et Δy sont des nombres assez petits pour que le point $M_1(x + \Delta x; y + \Delta y)$ appartienne à ce δ-voisinage de M.

Introduisons la fonction auxiliaire

$$\varphi(x) = f(x, y + \Delta y) - f(x, y);$$

l'expression A peut alors être traitée comme l'accroissement de la fonction $\varphi(x)$ d'une seule variable x, dérivable sur l'intervalle $[x, x + \Delta x]$:

$$A = \Delta\varphi = \varphi(x + \Delta x) - \varphi(x).$$

Le théorème de Lagrange appliqué à cette différence nous permet d'écrire

$$A = \Delta\varphi = \varphi'(x + \theta_1 \Delta x)\,\Delta x =$$
$$= [f'_x(x + \theta_1 \Delta x, y + \Delta y) - f'_x(x + \theta_1 \Delta x, y)]\Delta x, \quad 0 < \theta_1 < 1.$$

L'expression entre crochets peut être traitée comme l'accroissement de la fonction $f'_x(x + \theta_1 \Delta x, y)$ d'une seule variable y dérivable sur l'intervalle $[y, y + \Delta y]$. En appliquant le théorème de Lagrange (par rapport à y), on obtient

$$A = f''_{yx}(x + \theta_1 \Delta x, y + \theta_2 \Delta y)\,\Delta x\,\Delta y, \quad 0 < \theta_1, \theta_2, < 1. \quad (1)$$

D'autre part, si l'on introduit la fonction auxiliaire

$$\psi(y) = f(x + \Delta x, y) - f(x, y),$$

on obtient par analogie

$$A = \Delta\psi = \psi(y + \Delta y) - \psi(y),$$

et ensuite

$$A = f''_{xy}(x + \theta_4 \Delta x, \quad y + \theta_3 \Delta y)\,\Delta y \Delta x, \; 0 < \theta_3, \quad \theta_4 < 1. \quad (2)$$

En comparant (1) et (2), on trouve

$$f''_{yx}(x + \theta_1 \Delta x, y + \theta_2 \Delta y) = f''_{xy}(x + \theta_4 \Delta x, y + \theta_3 \Delta y).$$

En faisant tendre Δx et Δy vers 0 et en tenant compte de la continuité des dérivées partielles $f''_{yx}(x, y)$ et $f''_{xy}(x, y)$ en M, on trouve

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} f''_{yx}(x + \theta_1 \Delta x, y + \theta_2 \Delta y) = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} f''_{xy}(x + \theta_4 \Delta x, y + \theta_3 \Delta y)$$

ou

$$f''_{xy}(x, y) = f''_{yx}(x, y). \quad \blacksquare$$

On définit de façon analogue les dérivées partielles d'ordre trois, quatre et, plus généralement, d'ordre n et on démontre un théorème de type 12.5 sur l'égalité des dérivées mixtes de tout ordre.

2. Différentielles supérieures. Au § 4 nous avons introduit la notion de différentielle d'une fonction $z = f(M)$ dérivable en un point M et avons établi la formule

$$dz = f'_x(x, y)\,dx + f'_y(x, y)\,dy. \quad (3)$$

Appelons dz *différentielle première* ou *du premier ordre*. Par souci de commodité, les différentielles seront désignées aussi par le symbole δ.

Supposons que les fonctions $f'_x(x, y)$ et $f'_y(x, y)$ sont dérivables en M. Dans l'expression (3) de dz traitons dx et dy comme des constantes. La fonction dz est alors une fonction des seules variables x et y, dérivable en M et de différentielle

$$\delta(dz) = \delta[f'_x(x, y)\,dx + f'_y(x, y)\,dy] = \\ = [f'_x(x, y)\,dx + f'_y(x, y)\,dy]'_x\,\delta x + \\ + [f'_x(x, y)\,dx + f'_y(x, y)\,dy]'_y\,\delta y. \quad (4)$$

La différentielle $\delta(dz)$ de la différentielle dz au point M prise pour $\delta x = dx$, $\delta y = dy$, s'appelle *différentielle seconde* ou *d'ordre deux* de la fonction $z = f(M)$ en M et se note d^2z. La différentielle $\delta(d^2z)$ de d^2z prise pour $\delta x = dx$, $\delta y = dy$ s'appelle *différentielle troisième* ou *d'ordre trois* de la fonction $z = f(M)$ et se note d^3z. Plus généralement, la différentielle $\delta(d^{n-1}z)$ de la différentielle $d^{n-1}z$ prise pour $\delta x = dx$, $\delta y = dy$, s'appelle *différentielle n-ième* ou *d'ordre n* de la fonction $z = f(M)$ et se note d^nz.

Ainsi, pour la différentielle d'ordre n de la fonction $z = f(M)$ on a la formule

$$d^nz = \delta(d^{n-1}z)\Big|_{\substack{\delta x = dx \\ \delta y = dy}}.$$

Trouvons l'expression de la différentielle seconde à l'aide de la formule (4)

$$d^2z = \delta(dz)\Big|_{\substack{\delta x = dx \\ \delta y = dy}} = (f'_x\,dx + f'_y\,dy)'_x\,dx + (f'_x\,dx + f'_y\,dy)'_y\,dy = \\ = f''_{xx}(dx)^2 + f''_{xy}\,dx\,dy + f''_{yx}\,dy\,dx + f''_{yy}(dy)^2.$$

Si f''_{xy} et f''_{yx} sont continues, les termes $f''_{xy}\,dx\,dy$ et $f''_{yx}\,dy\,dx$ sont égaux d'après le théorème 12.5 et l'on a

$$d^2z = f''_{xx}(dx)^2 + 2f''_{xy}\,dx\,dy + f''_{yy}(dy)^2.$$

De façon analogue,

$$d^3z = f'''_{x^3}(dx)^3 + 3f'''_{x^2y}(dx)^2\,dy + 3f'''_{xy^2}\,dx\,(dy)^2 + f'''_{y^3}(dy)^3,$$

. . . . . .

$$d^nz = f^{(n)}_{x^n}(dx)^n + nf^{(n)}_{x^{n-1}y}(dx)^{n-1}\,dy + \ldots$$

$$\ldots + \frac{n(n-1)\ldots(n-k+1)}{k!}\,f^{(n)}_{x^{n-k}y^k}(dx)^{n-k}(dy)^k + \ldots + f^{(n)}_{y^n}(dy)^n.$$

La formule de d^nz rappelle le développement d'un binôme de degré n par la formule de Newton. C'est pourquoi l'expression de

$d^n z$ s'écrit sous la forme symbolique suivante plus facile à retenir :

$$d^n z = \left(\frac{\partial}{\partial x}\,dx + \frac{\partial}{\partial y}\,dy\right)^n f(x, y).$$

Exemples.

1. Calculer la différentielle seconde $d^2 z$ de la fonction $z = \text{Arc tg}\,\frac{x}{y}$. On a

$$f'_x = \frac{y}{x^2+y^2},\quad f'_y = -\frac{x}{x^2+y^2},\quad f''_{xy} = \frac{x^2-y^2}{(x^2+y^2)^2} = f''_{yx},$$

$$f''_{x^2} = -\frac{2xy}{(x^2+y^2)^2},\quad f''_{y^2} = \frac{2xy}{(x^2+y^2)^2}.$$

Donc

$$d^2 z = \frac{-2xy\,(dx)^2 + 2(x^2-y^2)\,dx\,dy + 2xy\,(dy)^2}{(x^2+y^2)^2}.$$

2. Calculer la différentielle troisième $d^3 z$ de la fonction $z = \sin x \cos y$. On a

$$f'_x = \cos x \cos y,\quad f'_y = -\sin x \sin y,\quad f''_{x^2} = -\sin x \cos y,$$
$$f''_{y^2} = -\sin x \cos y,\quad f''_{xy} = -\cos x \sin y,$$
$$f'''_{x^3} = -\cos x \cos y,\quad f'''_{y^3} = \sin x \sin y,$$
$$f'''_{x^2 y} = \sin x \sin y,\quad f'''_{xy^2} = -\cos x \cos y.$$

Donc

$$d^3 z = -\cos x \cos y\,(dx)^3 + 3 \sin x \sin y\,(dx)^2\,dy - \\ - 3 \cos x \cos y\,dx\,(dy)^2 + \sin x \sin y\,(dy)^3.$$

§ 7. Formule de Taylor pour les fonctions de deux variables

Par analogie à une fonction d'une variable, une fonction de deux variables peut être représentée par la somme d'un polynôme de degré n et d'un certain reste. Prouvons le théorème suivant.

Théorème 12.6. *Soit $z = f(x, y)$ une fonction continue avec ses dérivées partielles jusqu'à l'ordre $n+1$ dans un δ-voisinage d'un point $M(x, y)$ et soit $M_1(x + \Delta x;\ y + \Delta y)$ un point de ce voisinage. L'accroissement $\Delta f = f(M_1) - f(M)$ de cette fonction au point M peut alors être représenté sous la forme*

$$\Delta f = df(x, y) + \frac{d^2 f(x, y)}{2!} + \cdots + \frac{d^n f(x, y)}{n!} + \\ + \frac{d^{n+1} f(x+\theta\Delta x,\ y+\theta\Delta y)}{(n+1)!},\quad 0 < \theta < 1. \qquad (1)$$

La formule (1) s'appelle *formule de Taylor* pour la fonction $z = f(x, y)$.

Démonstration. Considérons la fonction auxiliaire

$$F(t) = f(x + t\Delta x,\ y + t\Delta y),$$

qui est une fonction composée de la variable indépendante t, $t \in [0, 1]$, $n + 1$ fois dérivable par rapport à t sur $[0; 1]$.

Une dérivation de $F(t)$ par rapport à t nous donne

$$F'(t) = f'_x(x + t\Delta x,\ y + t\Delta y)\,\Delta x + f'_y(x + t\Delta x,\ y + t\Delta y)\,\Delta y =$$

$$= \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right) f(x + t\Delta x,\ y + t\Delta y),$$

$$F''(t) = f''_{xx}(x + t\Delta x,\ y + t\Delta y)\,(\Delta x)^2 +$$

$$+ 2f''_{xy}(x + t\Delta x,\ y + t\Delta y)\,\Delta x\Delta y + f''_{yy}(x + t\Delta x,\ y + t\Delta y)\,(\Delta y)^2 =$$

$$= \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^2 f(x + t\Delta x,\ y + t\Delta y).$$

En raisonnant par récurrence, on obtient

$$F^{(n)}(t) = \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^n f(x + t\Delta x,\ y + t\Delta y),$$

$$F^{(n+1)}(t) = \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^{n+1} f(x + t\Delta x,\ y + t\Delta y).$$

En appliquant d'autre part la formule de Maclaurin (cf. chap. VI, § 3, n° 3) à la fonction $F(t)$ traitée comme une fonction de la variable t et en faisant $t = 1$, on trouve

$$F(1) = F(0) + \frac{F'(0)}{1!} + \frac{F''(0)}{2!} + \cdots + \frac{F^{(n)}(0)}{n!} + \frac{F^{(n+1)}(\theta)}{(n+1)!},$$

$$0 < \theta < 1. \qquad (2)$$

Mais

$$F(1) = f(x + \Delta x,\ y + \Delta y) = f(M_1),$$

$$F(0) = f(x,\ y) = f(M),$$

$$F'(0) = \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right) f(x,\ y) = \mathrm{d}f(x,\ y),$$

$$F''(0) = \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^2 f(x,\ y) = \mathrm{d}^2 f(x,\ y),$$

. . . . . .

$$F^{(n)}(0) = \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^n f(x,\ y) = \mathrm{d}^n f(x,\ y),$$

$$F^{(n+1)}(\theta) = \left(\frac{\partial}{\partial x}\Delta x + \frac{\partial}{\partial y}\Delta y\right)^{n+1} f(x + \theta\Delta x,\ y + \theta\Delta y) =$$

$$= \mathrm{d}^{n+1} f(x + \theta\Delta x,\ y + \theta\Delta y).$$

Ces égalités et la formule (2) entraînent

$$F(1)-F(0)=f(M_1)-f(M)=\Delta f=df(x, y)+\frac{d^2f(x, y)}{2!}+\ldots$$

$$\ldots+\frac{d^nf(x, y)}{n!}+\frac{d^{n+1}f(x+\theta\Delta x, y+\theta\Delta y)}{(n+1)!},\quad 0<\theta<1,$$

c'est-à-dire la formule (1). ■

La formule de Taylor d'une fonction de deux variables rappelle celle d'une fonction d'une variable. Mais en réalité, si l'on développe les expressions des différentielles de la fonction $f(x, y)$ dans la formule (1), on obtient une formule bien plus volumineuse et compliquée que pour une fonction d'une seule variable.

La formule de Taylor pour des fonctions d'un plus grand nombre de variables est de la même forme.

Remarque. Pour $n=0$, la formule (1) se transforme en la formule des accroissements finis de Lagrange pour une fonction de deux variables:

$$\Delta f=df(x+\theta\Delta x, y+\theta\Delta y)=$$
$$=f'_x(x+\theta\Delta x, y+\theta\Delta y)\,\Delta x+$$
$$+f'_y(x+\theta\Delta x, y+\theta\Delta y)\,\Delta y,\quad 0<\theta<1,$$

formule d'où il suit, en particulier, que si $f'_x=f'_y\equiv 0$, alors l'accroissement total de la fonction est identiquement nul et la fonction $f(x, y)$ est constante.

§ 8. Extrémums d'une fonction de deux variables

1. Définition de l'extrémum. Soit $z=f(x, y)$ une fonction définie sur un voisinage d'un point $M_0(x_0; y_0)$.

Définition. On dit qu'une fonction $z=f(x, y)$ présente en un point M_0 un *maximum* (resp. *minimum*) *local* s'il existe un voisinage du point M_0 sur lequel $f(x, y)\leqslant f(x_0, y_0)$ (resp. $f(x, y)\geqslant f(x_0, y_0)$) pour tout point $M(x; y)$.

Les points en lesquels une fonction présente un minimum ou un maximum local s'appelle *points d'extrémums locaux*. De la définition il résulte que si la fonction $z=f(x, y)$ présente un extrémum en M_0, son accroissement total $\Delta z=f(M)-f(M_0)$ en M_0 satisfait dans un voisinage de M_0 à l'une des conditions suivantes:

$$\Delta z\leqslant 0 \text{ en cas de maximum local,}$$
$$\Delta z\geqslant 0 \text{ en cas de minimum local.}$$

Inversement, si l'une de ces inégalités a lieu au voisinage d'un point M_0, la fonction présente un extrémum local en M_0.

2. Conditions nécessaires d'extrémum.

Théorème 12.7. *Si une fonction $f(x, y)$ possède des dérivées partielles premières en un point $M_0(x_0; y_0)$ où elle présente un extrémum, ces dérivées sont nulles en M_0, c'est-à-dire que*

$$f'_x(x_0, y_0) = f'_y(x_0, y_0) = 0. \quad (1)$$

Démonstration. Prouvons, par exemple, que $f'_x(x_0, y_0) = 0$. Considérons à cet effet les seuls points du voisinage de M_0 pour lesquels $y = y_0$. On obtient ainsi une fonction $f(x, y_0)$ d'une seule variable x qui au point $x = x_0$ est dérivable et présente un extrémum. Donc $f'_x(x_0, y_0) = 0$ d'après la condition nécessaire d'extrémum d'une fonction d'une variable. Ce qu'on voulait.

En envisageant la fonction $f(x_0, y)$ de la variable y, on établit par analogie que $f'_y(x_0, y_0) = 0$. ■

La condition (1) n'est pas une condition suffisante d'extrémum. Par exemple, les dérivées partielles de la fonction $z = x^2 - y^2$ sont nulles au point $(0; 0)$, néanmoins cette fonction ne présente pas d'extrémum en $(0; 0)$, car elle y est nulle et ne garde son signe dans aucun voisinage de ce point: si $x = 0$, alors $z < 0$, si $y = 0$, alors $z > 0$. La fonction $z = x^2 - y^2$ est représentée graphiquement par un paraboloïde hyperbolique (cf. fig. 160).

Ainsi, la condition (1) n'est qu'une condition nécessaire d'extrémum. Aux points où elle est réalisée on dira par analogie avec le cas d'une variable que la fonction présente un *éventuel extrémum*. Ces points sont appelés aussi stationnaires.

3. Conditions suffisantes d'extrémum.

Théorème 12.8. *Supposons qu'une fonction $f(x, y)$ possède des dérivées partielles secondes continues sur un voisinage d'un point stationnaire $M_0(x_0; y_0)$. Posons*

$$\Delta = \begin{vmatrix} f''_{xx}(x_0, y_0) & f''_{xy}(x_0, y_0) \\ f''_{xy}(x_0, y_0) & f''_{yy}(x_0, y_0) \end{vmatrix}.$$

Si

a) $\Delta > 0$, *la fonction $f(x, y)$ présente un extrémum local en M_0. Cet extrémum est un maximum local si $f''_{xx}(x_0, y_0) < 0$ et un minimum local si $f''_{xx}(x_0, y_0) > 0$;*

b) $\Delta < 0$, *la fonction $f(x, y)$ ne présente pas d'extrémum en M_0.*

Démonstration. a) Soit $\Delta > 0$. Introduisons les notations suivantes: $f''_{xx}(x_0, y_0) = A$, $f''_{xy}(x_0, y_0) = B$ et $f''_{yy}(x_0, y_0) = C$. Par hypothèse, $f'_x(x_0, y_0) = f'_y(x_0, y_0) = 0$, $A > 0$ (ou $A < 0$). D'après la formule de Taylor (1) du § 7 prise pour $n = 1$, l'accroissement total de la fonction $f(x, y)$ au point M_0

peut être mis sous la forme

$$\Delta f = \frac{1}{2!}\,[A'(\Delta x)^2 + 2B'\Delta x\,\Delta y + C'(\Delta y)^2], \qquad (1)$$

où $A' = f''_{xx}(x_0 + \theta\Delta x,\ y_0 + \theta\Delta y)$, $B' = f''_{xy}(x_0 + \theta\Delta x,\ y_0 + \theta\Delta y)$, $C' = f''_{yy}(x_0 + \theta\Delta x,\ y_0 + \theta\Delta y)$, $0 < \theta < 1$. Les dérivées partielles secondes étant continues en M_0, on a

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} A' = f''_{xx}(x_0,\ y_0) = A > 0 \text{ (ou } A < 0),$$

et

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} (A'C' - B'^2) = f''_{xx}(x_0,\ y_0)\, f''_{yy}(x_0,\ y_0) - [f''_{xy}(x_0,\ y_0)]^2 = \Delta > 0.$$

Pour Δy et Δx assez petits, on a donc

$$A' > 0 \quad (\text{ou } A' < 0), \qquad A'C' - B'^2 = \Delta' > 0.$$

Comme $A' \neq 0$, la relation (1) peut être mise sous la forme

$$\Delta f = \frac{1}{2!} \cdot \frac{1}{A'}\,[A'^2(\Delta x)^2 + 2A'B'\Delta x\Delta y + A'C'(\Delta y)^2],$$

ou, si l'on met le trinôme entre crochets sous sa forme canonique,

$$\Delta f = \frac{1}{2!}\,\frac{1}{A'}\,[(A'\Delta x + B'\Delta y)^2 + (A'C' - B'^2)(\Delta y)^2].$$

L'expression entre crochets est positive, donc si $A' > 0$ ($f''_{xx}(x_0,\ y_0) > 0$), alors $\Delta f \geqslant 0$ et par suite on a un minimum local en M_0 ; si $A' < 0$ ($f''_{xx}(x_0,\ y_0) < 0$) alors $\Delta f \leqslant 0$ et l'on a un maximum local en M_0. C.Q.F.D.

c) Supposons maintenant que $\Delta = AC - B^2 < 0$ et, comme précédemment, que $A = f''_{xx}(x_0,\ y_0)$, $B = f''_{xy}(x_0,\ y_0)$, $C = f''_{yy}(x_0,\ y_0)$. Considérons le trinôme

$$A + 2Bx + Cx^2.$$

Puisque $B^2 - AC > 0$, on peut exhiber deux nombres x_1 et x_2 tels que

$$A + 2Bx_1 + Cx_1^2 > 0, \qquad A + 2Bx_2 + Cx_2^2 < 0.$$

Comme en a), mettons l'accroissement total de la fonction $f(x, y)$ en M_0 sous la forme (1). Les dérivées partielles secondes étant continues, on a

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} (A' + 2B'x_1 + C'x_1^2) = A + 2Bx_1 + Cx_1^2 > 0.$$

Il existe donc un δ-voisinage du point M_0 tel que si M ($x_0 + \Delta x$; $y_0 + \Delta y$) lui appartient, alors

$$A' + 2B'x_1 + C'x_1^2 > 0. \qquad (2)$$

Considérons maintenant un δ'-voisinage du point M_0 tel que $\delta' \leqslant \delta$. On peut choisir un nombre $t > 0$ assez petit pour que le point $M_1 (x + t;\ y + tx_1)$ appartienne au δ'-voisinage de M_0. En posant $\Delta x = t$ et $\Delta y = tx_1$ dans (1), on obtient en vertu de (2)

$$\Delta f = f(x_0 + \Delta x,\ y_0 + \Delta y) - f(x_0,\ y_0) = \\ = \frac{1}{2!} t^2 [A' + 2B'x_1 + C'x_1^2] > 0.$$

En raisonnant de même relativement à x_2, on trouve que le δ'-voisinage du point M_0 contient un point $M_2 (x + \Delta x;\ y + \Delta y)$ tel que

$$\Delta f = f(x_0 + \Delta x,\ y_0 + \Delta y) - f(x_0,\ y_0) < 0,$$

c'est-à-dire que l'accroissement de la fonction $f(x, y)$ ne garde son signe dans aucun voisinage si petit soit-il de M_0, donc il n'y a pas d'extrémum en M_0. ■

Remarque. Si $\Delta = 0$, la fonction $f(x, y)$ peut présenter ou non un extrémum en un point stationnaire.

Exemples. Trouver l'extrémum des fonctions:

1. $z = x^2 + xy + y^2 - 2x - 3y$. On a

$$f'_x = 2x + y - 2,\ f'_y = x + 2y - 3.$$

Cherchons les points stationnaires. Résolvons à cet effet le système d'équations

$$\begin{cases} 2x + y - 2 = 0, \\ x + 2y - 3 = 0. \end{cases}$$

La solution est $x = 1/3$, $y = 4/3$. Donc $M_0 (1/3;\ 4/3)$ est un point stationnaire.

Par ailleurs, $f''_{xx} = 2$, $f''_{xy} = 1$, $f''_{yy} = 2$, $\Delta = 2 \cdot 2 - 1 = 3$. Comme $\Delta = 3 > 0$ et $f''_{xx} = 2 > 0$, la fonction étudiée présente un minimum en $M_0 (1/3;\ 4/3)$.

2. $z = x^2 - y^2$. On a $f'_x = 2x$, $f'_y = -2y$. En résolvant le système d'équations $2x = 0$, $-2y = 0$, on trouve que $M_0 (0;\ 0)$ est un point stationnaire. Puisque $f''_{xx} = 2$, $f''_{xy} = 0$, $f''_{yy} = -2$ et donc que $\Delta = 2 \cdot (-2) - 0 = -4 < 0$, la fonction z ne présente pas d'extrémum en $M_0 (0;\ 0)$.

3. $z = x^4 + y^4$. On a $f'_x = 4x^3$, $f'_y = 4y^3$, $f''_{xx} = 12x^2$, $f''_{xy} = 0$, $f''_{yy} = 12y^2$. En résolvant le système d'équations $4x^3 = 0$, $4y^3 = 0$, on trouve que le point $M_0 (0;\ 0)$ est stationnaire. D'autre part, $f''_{xx} (0, 0) = 0$, $f''_{yy} (0, 0) = 0$, donc $\Delta = 0$. D'après la remarque ci-dessus, la fonction z peut présenter ou non un extrémum en M_0. Ici on a un extrémum, puisque $z > 0$ en tous les points

sauf en M_0 où elle est nulle. Donc la fonction z présente un minimum en M_0.

4. $z = x^3 + y^3$. On a $f'_x = 3x^2$, $f'_y = 3y^2$, $f''_{xx} = 6x$, $f''_{xy} = 0$, $f''_{yy} = 6y$. En résolvant le système d'équations $3x^2 = 0$, $3y^2 = 0$, on trouve que le point M_0 (0 ; 0) est un point stationnaire. D'autre part, f''_{xx} (0, 0) $= 0$, f''_{yy} (0, 0) $= 0$, donc $\Delta = 0$. La fonction z ne présente pas d'extrémum ici. En effet, z (0, 0) $= 0$, z (x, 0) $= x^3$, d'où $z > 0$ pour $x > 0$ et $z < 0$ pour $x < 0$, c'est-à-dire que dans tout voisinage du point M_0, cette fonction prend des valeurs aussi bien supérieures qu'inférieures à z (0, 0).

§ 9. Méthode des moindres carrés

Dans les recherches on est souvent conduit à utiliser des formules empiriques. L'une des meilleures méthodes d'établissement de ces formules est la *méthode des moindres carrés.*

Soit à établir une relation empirique entre deux variables x et y. Par exemple, entre la température et l'allongement linéaire d'une barre métallique. Les résultats des mesures sont consignés dans le tableau suivant :

x	x_1	x_2	...	x_i	...	x_n
y	y_1	y_2	...	y_i	...	y_n

Etablissons maintenant la forme de la fonction $y = f(x)$ par le nuage de points $(x_1 ; y_1)$, $(x_2 ; y_2)$, . . ., $(x_n ; y_n)$ de la figure 166.

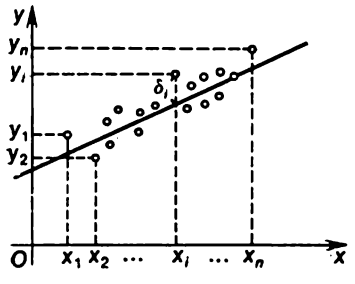


Fig. 166

Ce nuage nous suggère une dépendance linéaire entre x et y de la forme

$$y = ax + b \qquad (1)$$

qui est l'équation d'une droite appelée *droite de régression de y en x.*

Les points du nuage n'étant pas alignés, la formule (1) est une formule approchée. Donc en portant les coordonnées de ces points dans l'expression $y - (ax + b)$, on obtient les égalités $y_1 - (ax_1 + b) = \delta_1$, $y_2 - (ax_2 + b) = \delta_2$, . . ., $y_n - (ax_n + b) = \delta_n$, où δ_1, δ_2, . . ., δ_n sont des nombres appelés *erreurs.*

On se propose de choisir les coefficients a et b de telle sorte que ces erreurs soient minimales en valeur absolue. Utilisons à cet effet la méthode des moindres carrés. Soit la somme des carrés

des erreurs

$$S(a,\ b)=\sum_{i=1}^{n}[y_i-(ax_i+b)]^2=\sum_{i=1}^{n}\delta_i^2,$$

où x_i et y_i sont des nombres donnés, a et b des coefficients inconnus qu'il faut déterminer en minimisant $S(a, b)$.

Le problème se ramène donc à la détermination des valeurs de a et b minimisant la fonction $S(a, b)$. On a

$$\frac{\partial S}{\partial a}=-2\sum_{i=1}^{n}[y_i-(ax_i+b)]\,x_i,\qquad \frac{\partial S}{\partial b}=-2\sum_{i=1}^{n}[y_i-(ax_i+b)].$$

En égalant ces dérivées partielles à 0, on obtient un système linéaire de deux équations à deux inconnues a et b:

$$\begin{cases}\sum\limits_{i=1}^{n}y_ix_i=a\sum\limits_{i=1}^{n}x_i^2+b\sum\limits_{i=1}^{n}x_i,\\ \sum\limits_{i=1}^{n}y_i=a\sum\limits_{i=1}^{n}x_i+bn.\end{cases}\tag{2}$$

Le système (2) s'appelle *système normal de la méthode des moindres carrés*. Sa résolution nous donne les valeurs de a et b cherchées, donc l'équation (1) de la droite de régression.

On établit sans peine que la fonction $S(a, b)$ présente un minimum au point $M(a, b)$ à l'aide des dérivées partielles secondes. On a

$$\frac{\partial^2 S}{\partial a^2}=2\sum_{i=1}^{n}x_i^2,\qquad \frac{\partial^2 S}{\partial a\,\partial b}=2\sum_{i=1}^{n}x_i,\qquad \frac{\partial^2 S}{\partial b^2}=2n.$$

Donc

$$\Delta=\frac{\partial^2 S}{\partial a^2}\,\frac{\partial^2 S}{\partial b^2}-\left(\frac{\partial^2 S}{\partial a\,\partial b}\right)^2=4n\sum_{i=1}^{n}x_i^2-\left(2\sum_{i=1}^{n}x_i\right)^2.$$

Cette expression peut être mise sous la forme $\Delta=2\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}(x_i-x_j)^2$, d'où il s'ensuit que $\Delta>0$. La fonction $S(a, b)$ présente un minimum en $M(a; b)$ puisque $\frac{\partial^2 S}{\partial a^2}>0$.

Exemple. On donne le tableau suivant:

x	−2	0	1	2	4
y	0,5	1	1,5	2	3

Cherchons l'équation de la droite de régression de y en x.

Pour former le système (2), calculons préalablement les quantités suivantes :

$$\sum_{i=1}^{5} y_i x_i = 16{,}5\,; \quad \sum_{i=1}^{5} x_i^2 = 25\,; \quad \sum_{i=1}^{5} x_i = 5\,; \quad \sum_{i=1}^{5} y_i = 8.$$

Le système (2) s'écrit alors

$$\begin{cases} 25a + 5b = 16{,}5, \\ 5a + 5b = 8. \end{cases}$$

Sa résolution nous donne $a = 0{,}425$, $b = 1{,}175$. L'équation de la droite de régression est donc $y = 0{,}425\ x + 1{,}175$.

CHAPITRE XIII

INTÉGRATION

Dans ce chapitre nous étudions les principaux problèmes d'intégration des fonctions de deux variables. Les définitions et résultats acquis peuvent être généralisés aux fonctions de trois et de plus de trois variables.

§ 1. Intégrales doubles

L'intégrale double est l'extension de la notion d'intégrale définie à une fonction de deux variables.

1) Définition et conditions d'existence de l'intégrale double. Soient G un domaine fermé borné, $z = f(x, y)$ une fonction définie et bornée sur ce domaine.

On admet que la frontière du domaine G est composée d'un nombre fini de courbes définies par des équations données de la forme $y = f(x)$ ou $x = \varphi(y)$, où $f(x)$ et $\varphi(y)$ sont des fonctions continues. Comme exemples de tels domaines citons un polygone fermé dont la frontière est une ligne polygonale composée de segments de droite d'équations $y = kx + b$ ou $x = a$; le domaine borné par une ellipse (la frontière est composée ici des deux courbes $y = \pm \frac{b}{a}\sqrt{a^2 - x^2}$), etc.

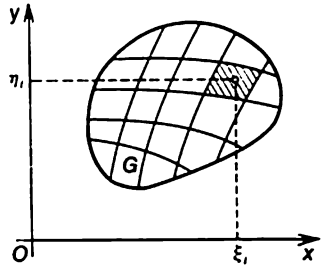


Fig. 167

Divisons le domaine G en n parties G_i ne possédant pas de points intérieurs communs, d'aires Δs_i ($i = 1, 2, \ldots, n$) (fig. 167). Dans chaque partie G_i choisissons arbitrairement un point $(\xi_i ; \eta_i)$ et formons la somme

$$\sigma = \sum_{i=1}^{n} f(\xi_i ; \eta_i)\, \Delta s_i, \qquad (1)$$

que nous appellerons *somme intégrale* de la fonction $f(x, y)$ dans le domaine G. Appelons *diamètre* d (G) du domaine G la plus

grande des distances de deux points frontières de G. Désignons par λ le plus grand des diamètres des domaines partiels G_i ($\lambda = \max_{1 \leqslant i \leqslant n} d(G_i)$).

Définition. La limite I *), si elle existe, de la somme intégrale (1) lorsque $\lambda \to 0$ s'appelle *intégrale double* de la fonction $f(x, y)$ étendue au domaine G et se représente par l'une des notations suivantes:

$$I = \int_G \int f(x, y)\, ds = \int_G \int f(x, y)\, dx\, dy.$$

La fonction $f(x, y)$ est dite *intégrable* sur le domaine G; le domaine G s'appelle *domaine d'intégration*; x et y, *variables d'intégration*; ds (ou $dxdy$) *élément d'aire* ou *aire élémentaire.*

En définissant l'intégrale double, nous avons admis que la fonction $f(x, y)$ était bornée. Comme dans le cas d'une fonction d'une seule variable, cette condition est une condition nécessaire d'intégrabilité. Mais elle n'est pas suffisante, c'est-à-dire qu'il existe des fonctions bornées qui ne sont pas intégrables. Comme exemple citons la fonction $f(x, y)$ définie sur le carré $\{(x;y) \mid 0 \leqslant x \leqslant 1;\ 0 \leqslant y \leqslant 1\}$ de la manière suivante:

$$f(x, y) = \begin{cases} 1 \text{ si } x \text{ et } y \text{ sont rationnels,} \\ 0 \text{ s'ils sont irrationnels.} \end{cases}$$

La non-intégrabilité de cette fonction résulte directement de la définition de l'intégrale double.

Pour établir les conditions suffisantes d'intégrabilité, il est commode de se servir, comme dans le cas d'une seule variable, de la théorie des sommes de Darboux qui se généralise intégralement à l'intégrale double **). Le théorème suivant se prouve par analogie au théorème correspondant de l'intégrale définie.

Théorème 13.1. *Une fonction $f(x, y)$ continue sur un domaine fermé borné G est intégrable sur ce domaine.*

Signalons que l'intégrale double n'existe que pour les seules fonctions continues. Nous avons un théorème plus général.

Théorème 13.2. *Une fonction $f(x, y)$, bornée sur un domaine fermé borné G et partout continue dans ce domaine sauf aux points*

*) Le nombre I s'appelle *limite de la somme intégrale* (1) lorsque $\lambda \to 0$ si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour $\lambda < \delta$ l'on ait $|\sigma - I| < \varepsilon$ indépendamment du choix des points $(\xi_i; \eta_i)$.

**) On démontre, en particulier, que si $f(x, y)$ est intégrable sur le domaine G, la limite des sommes de Darboux aussi bien inférieures que supérieures est égale à $\int_G \int f(x, y)\, ds$ lorsque $\lambda \to 0$.

d'un nombre fini de courbes continues d'équations $y = f(x)$ *ou* $x = \varphi(y)$, *est intégrable sur ce domaine.*

2. Interprétation géométrique de l'intégrale double. Soit P le solide (fig. 168) limité par une surface d'équation $z = f(x, y)$, le plan xOy et le cylindre de section droite G. Le solide P s'appelle *cylindre curviligne.*

De même que le calcul de l'aire du trapèze curviligne conduit à l'interprétation géométrique de l'intégrale définie, de même le calcul du volume de P conduit à l'interprétation géométrique de l'intégrale double.

En effet, la somme intégrale (1) est ici la somme des volumes de cylindres droits d'aires de base Δs_i et de hauteurs $f(\xi_i, \eta_i)$, qui peut être adoptée pour valeur approchée du volume de P :

$$v_P \approx \sum_{i=1}^{n} f(\xi_i, \eta_i)\,\Delta s_i.$$

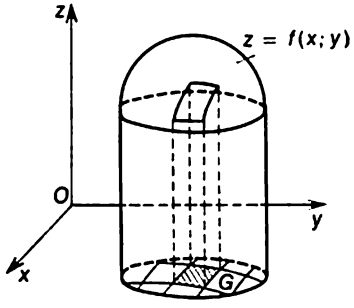


Fig. 168

Cette approximation est d'autant plus exacte que la subdivision du domaine G est fine. Si $\lambda \to 0$, cette égalité approchée se transforme en l'égalité rigoureuse :

$$v_P = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i, \eta_i)\,\Delta s_i.$$

La fonction $f(x, y)$ étant intégrable, sa somme intégrale admet une limite qui est égale à l'intégrale double de $f(x, y)$ étendue au domaine G. Donc

$$v_P = \int\int_G f(x, y)\,dx\,dy.$$

On voit donc que l'intégrale double d'une fonction continue est égale au volume d'un cylindre curviligne.

Remarque. Si l'on pose $f(x, y) \equiv 1$ sur le domaine G, la définition de l'intégrale double nous donne l'aire s du domaine G sous forme d'une intégrale double :

$$\int\int_G 1 \cdot dx\,dy = \lim_{\lambda \to 0} \sum_{i=1}^{n} 1 \cdot \Delta s_i = \lim_{\lambda \to 0} s = s.$$

3. Propriétés de l'intégrale double. Les propriétés fondamentales de l'intégrale double sont identiques aux propriétés corres-

pondantes de l'intégrale définie. Nous les formulerons donc sans les prouver.

1° Si k est une constante et $f(x, y)$ une fonction intégrable sur un domaine G, la fonction $kf(x, y)$ est aussi intégrable sur G et

$$\int_G \int kf(x, y)\,dx\,dy = k \int_G \int f(x, y)\,dx\,dy,$$

autrement dit on peut faire sortir une constante du signe d'intégration.

2° Si $f(x, y)$ et $g(x, y)$ sont des fonctions intégrables sur G, il en est de même de leur somme algébrique et

$$\int_G \int [f(x, y) \pm g(x, y)]\,dx\,dy = \int_G \int f(x, y)\,dx\,dy \pm \pm \int_G \int g(x, y)\,dx\,dy.$$

3° Une fonction $f(x, y)$ intégrable sur des domaines G_1 et G_2 sans points intérieurs communs est intégrable sur la réunion G de ces domaines et

$$\int_G \int f(x, y)\,dx\,dy = \int_{G_1} \int f(x, y)\,dx\,dy + \int_{G_2} \int f(x, y)\,dx\,dy.$$

4° **Théorème de la moyenne.** *Si une fonction $f(x, y)$ est intégrable sur un domaine G, il existe un point $(\xi_i ; \eta_i) \in G$ tel que*

$$\int_G \int f(x, y)\,dx\,dy = f(\xi_i ; \eta_i)\, s,$$

où s est l'aire limitée par G.

Maintenant que nous avons défini l'intégrale double, étudié ses propriétés fondamentales, ses conditions d'existence et son interprétation géométrique, passons à ses méthodes de calcul.

§ 2. Réduction de l'intégrale double à l'intégrale itérée

1. Cas d'un domaine rectangulaire. Etudions tout d'abord l'intégrale double sur un rectangle D de côtés parallèles aux axes de coordonnées.

Théorème 13.3. *Soit $f(x, y)$ une fonction définie sur un rectangle $D = \{(x ; y) \mid a \leqslant x \leqslant b ; c \leqslant y \leqslant d\}$. Supposons qu'existe l'intégrale double*

$$\int_D \int f(x, y)\,dx\,dy. \qquad (1)$$

Supposons, par ailleurs, que pour tout $x \in [a, b]$ *existe l'intégrale définie*

$$I(x) = \int_c^d f(x, y)\,dy. \tag{2}$$

Alors existe l'intégrale

$$\int_a^b I(x)\,dx = \int_a^b dx \int_c^d f(x, y)\,dy$$

(*appelée* intégrale itérée) *et*

$$\iint_D f(x, y)\,dx\,dy = \int_a^b dx \int_c^b f(x, y)\,dy. \tag{3}$$

Démonstration. Divisons le rectangle D par les points $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ et $c = y_0 < y_1 < y_2 < \ldots < y_k = d$ en nk éléments $D_{ij} = \{(x;\ y) \mid x_{i-1} \leqslant x \leqslant x_i;\ y_{j-1} \leqslant y \leqslant y_j\}$. Posons $\Delta x_i = x_i - x_{i-1}$, $\Delta y_j = y_j - y_{j-1}$ et désignons par m_{ij} et M_{ij} respectivement l'infimum et le suprémum de la fonction $f(x, y)$ sur l'élément D_{ij} (fig. 169). On a alors sur cet élément

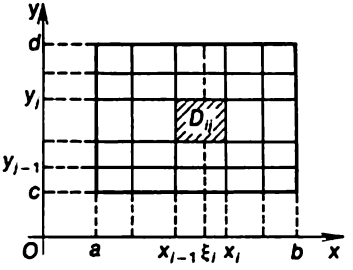


Fig. 169

$$m_{ij} \leqslant f(x, y) \leqslant M_{ij}. \tag{4}$$

Posons dans cette double inégalité $x = \xi_i$, où ξ_i est un point arbitraire de l'intervalle $[x_{i-1}, x_i]$ et intégrons ensuite (4) par rapport à y entre y_{j-1} et y_j. On obtient

$$m_{ij}\Delta y_j \leqslant \int_{y_{j-1}}^{y_j} f(\xi_i, y)\,dy \leqslant M_{ij}\Delta y_j. \tag{5}$$

En sommant (5) sur tous les j entre 1 et k et en utilisant la notation (2), on obtient

$$\sum_{j=1}^{k} m_{ij}\Delta y_j \leqslant I(\xi_i) \leqslant \sum_{j=1}^{k} M_{ij}\Delta y_j. \tag{6}$$

D'autre part, en multipliant (6) par Δx_i et en sommant sur tous les i entre 1 et n, on trouve

$$\sum_{i=1}^{n}\sum_{j=1}^{k} m_{ij}\Delta x_i \Delta y_j \leqslant \sum_{i=1}^{n} I(\xi_i)\,\Delta x_i \leqslant \sum_{i=1}^{n}\sum_{j=1}^{k} M_{ij}\Delta x_i \Delta y_j. \tag{7}$$

Supposons que le plus grand des diamètres des éléments D_{ij} tend vers 0 ($\lambda \to 0$). La plus grande longueur Δx_i tend aussi vers 0. Les membres extrêmes de (7) qui sont les sommes inférieure et supérieure de Darboux tendent alors vers l'intégrale double (1) (cf. note de la page 344). Donc le membre moyen de (7) admet une limite qui est égale à la même intégrale double. Or, d'après la définition de l'intégrale, cette limite est égale à

$$\int_a^b I(x)\,dx = \int_a^b dx \int_c^d f(x, y)\,dy.$$

On a ainsi prouvé l'existence de l'intégrale itérée et l'égalite (3). ■

Remarque. Si dans le théorème 13.3 on intervertit x et y, on démontre l'existence de l'intégrale itérée

$$\int_c^d I(y)\,dy = \int_c^d dy \int_a^b f(x, y)\,dx$$

et l'égalité

$$\iint_D f(x, y)\,dx\,dy = \int_c^d dy \int_a^b f(x, y)\,dx. \qquad (8)$$

L'intégrale double se ramène à l'intégrale itérée grâce aux formules (3) et (8). Ainsi, par exemple, dans la formule (8) on intègre par rapport à x à y constant et ensuite par rapport à y, c'est-à-dire qu'on prend successivement deux intégrales définies.

Exemple. Calculer $\iint_D xy\,dx\,dy$, où $D = \{(x; y) \mid 1 \leqslant x \leqslant 2;\ 1 \leqslant y \leqslant 2\}$.

Solution. On a

$$\iint_D xy\,dx\,dy = \int_1^2 dy \int_1^2 xy\,dx = \int_1^2 \left[y\frac{x^2}{2}\right]_1^2 dy =$$

$$= \int_1^2 \left[2y - \frac{1}{2}y\right] dy = \frac{3}{4}y^2\Big|_1^2 = \frac{9}{4}.$$

2. Cas d'un domaine curviligne.

Théorème 13.4. *Soit $z = f(x, y)$ une fonction définie sur un domaine $G = \{(x; y) \mid a \leqslant x \leqslant b;\ y_1(x) \leqslant y \leqslant y_2(x)\}$, où $y_1(x)$ et $y_2(x)$ sont des fonctions continues. Si existent l'intégrale double*

$$\iint_G f(x, y)\,dx\,dy$$

et, pour tout $x \in [a, b]$, *l'intégrale définie*

$$I(x) = \int_{y_1(x)}^{y_2(x)} f(x, y)\, dy,$$

alors existe l'intégrale itérée

$$\int_a^b I(x)\, dx = \int_a^b dx \int_{y_1(x)}^{y_2(x)} f(x, y)\, dy$$

et l'on a l'égalité

$$\int\int_G f(x, y)\, dx\, dy = \int_a^b dx \int_{y_1(x)}^{y_2(x)} f(x, y)\, dy. \tag{9}$$

Démonstration. Posons $c = \min\limits_{[a, b]} y_1(x)$, $d = \max\limits_{[a, b]} y_2(x)$ et comprenons le domaine G dans le rectangle $D = \{(x; y) \mid a \leqslant x \leqslant b;\ c \leqslant y \leqslant d\}$ (fig. 170). Considérons sur ce rectangle la fonction auxiliaire

$$F(x, y) = \begin{cases} f(x, y) & \text{dans } G, \\ 0 & \text{dans } D \setminus G. \end{cases}$$

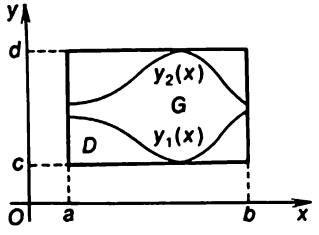


Fig. 170

Cette fonction satisfait à toutes les conditions du théorème précédent. En effet, elle est intégrable sur G, puisqu'elle est confondue avec $f(x, y)$, et intégrable sur $D \setminus G$ où elle est nulle. Donc d'après la propriété 3° du § 1, elle est intégrable sur le rectangle D tout entier. De plus

$$\int\int_G F(x, y)\, dx\, dy = \int\int_G f(x, y)\, dx\, dy$$

$$\text{et} \int\int_{D \setminus G} F(x, y)\, dx\, dy = 0,$$

d'où

$$\int\int_D F(x, y)\, dx\, dy = \int\int_G f(x, y)\, dx\, dy. \tag{10}$$

D'autre part, pour tout $x \in [a, b]$ existe l'intégrale

$$\int_c^d F(x, y)\, dy = \int_c^{y_1(x)} F(x, y)\, dy + \int_{y_1(x)}^{y_2(x)} F(x, y)\, dy + \int_{y_2(x)}^d F(x, y)\, dy,$$

puisque chacune des trois intégrales du second membre existe. En effet, les segments $[c, y_1(x)]$ et $[y_2(x), d]$ sont extérieurs au domaine G et $F(x, y)$ est nulle sur eux, donc la première et la troisième intégrale sont nulles, quant à la deuxième, elle existe par hypothèse, car $F(x, y) = f(x, y)$ sur l'intervalle $[y_1(x), y_2(x)]$. Donc

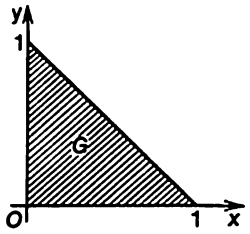


Fig. 171

$$\int_c^d F(x, y)\,dy = \int_{y_1(x)}^{y_2(x)} f(x, y)\,dy. \tag{11}$$

La fonction $F(x, y)$ satisfaisant ainsi à toutes les conditions du théorème 13.3, son intégrale double sur D peut être ramenée à une intégrale itérée:

$$\iint_D F(x, y)\,dx\,dy = \int_a^b dx \int_c^d F(x, y)\,dy.$$

De là et des égalités (10) et (11) on déduit que

$$\iint_G f(x, y)\,dx\,dy = \int_a^b dx \int_{y_1(x)}^{y_2(x)} f(x, y)\,dy,$$

c'est-à-dire la formule (9). ■

Remarque 1. Si dans le théorème 13.4 on intervertit x et y, on établit l'existence de l'intégrale itérée

$$\int_c^d I(y)\,dy = \int_c^d dy \int_{x_1(y)}^{x_2(y)} f(x, y)\,dx$$

et l'égalité

$$\iint_G f(x, y)\,dx\,dy = \int_c^d dy \int_{x_1(y)}^{x_2(y)} f(x, y)\,dx. \tag{12}$$

Exemple. Calculer l'intégrale $\iint_G (x^2 + xy + 2y^2)\,dx\,dy$ sur le domaine $G = \{(x; y) \mid 0 \leqslant x \leqslant 1;\ 0 \leqslant y \leqslant 1 - x\}$.

Solution. Le domaine G est le triangle limité par les axes de coordonnées et la droite $y = -x + 1$ (fig. 171). Donc

$y_1(x) = 0$, $y_2(x) = 1 - x$. La formule (9) nous donne

$$\iint_G (x^2 + xy + 2y^2)\,dx\,dy = \int_0^1 dx \int_0^{1-x} (x^2 + xy + 2y^2)\,dy =$$

$$= \int_0^1 \left[x^2(1-x) + \frac{x(1-x)^2}{2} + \frac{2(1-x)^3}{3}\right] dx = \frac{7}{24}.$$

On peut calculer cette intégrale à l'aide de la formule (12) si l'on intervertit x et y dans G. Le triangle est alors défini par les doubles inégalités $0 \leqslant y \leqslant 1$, $0 \leqslant x \leqslant 1 - y$, d'où $x_1(y) = 0$, $x_2(y) = 1 - y$ et l'on vérifie immédiatement que l'intégrale

$$\iint_G (x^2 + xy + 2y^2)\,dx\,dy = \int_0^1 dy \int_0^{1-y} (x^2 + xy + 2y^2)dx$$

prend la même valeur.

Remarque 2. Si le domaine G ne satisfait pas aux conditions du théorème 13.4 (par exemple il est coupé en plus de deux points par des droites verticales ou horizontales), il faut le partager en parties vérifiant chacune ces conditions et ramener ensuite à une intégrale itérée chaque intégrale double étendue à ces parties.

§ 3. Changement de variables dans l'intégrale double

Soit $f(x, y)$ une fonction continue sur un domaine fermé borné G. Alors existe l'intégrale double

$$\iint_G f(x, y)\,dx\,dy. \tag{1}$$

Considérons le changement de variables

$$x = x(u, v), \quad y = y(u, v). \tag{2}$$

On admettra que u et v se définissent de façon unique à partir de (2) :

$$u = u(x, y), \quad v = v(x, y). \tag{3}$$

Les formules (3) associent à chaque point $M(x; y)$ du domaine G un point $M^*(u; v)$ d'un plan rapporté aux coordonnées rectangulaires u et v. Supposons que les points $M^*(u; v)$ forment un domaine fermé borné G^*. Les formules (2) s'appellent *formules de transformation des coordonnées*, les formules (3), *formules de transformation réciproque*.

Sous les hypothèses adoptées on démontre que si les fonctions (2) admettent sur G^* des dérivées partielles premières conti-

nues et si le déterminant

$$\frac{D(x, y)}{D(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \tag{4}$$

est différent de 0 sur G^*, l'intégrale (1) est justiciable de la formule de changement des variables

$$\int\int_G f(x, y)\,dx\,dy = \int\int_{G^*} f[x(u, v), y(u, v)] \left| \frac{D(x, y)}{D(u, v)} \right| du\,dv. \tag{5}$$

Le déterminant (4) s'appelle *déterminant fonctionnel* ou *jacobien* (du nom du mathématicien allemand Jacobi) des fonctions $x = x(u, v)$, $y = y(u, v)$ par rapport aux variables u et v.

On peut résumer ce qui précède par le théorème suivant.

Théorème 13.5. *Si la transformation* (2) *envoie bijectivement le domaine fermé borné G sur un domaine fermé borné G^*; si les fonctions* (2) *possèdent des dérivées partielles premières continues sur G^* et un jacobien* (4) *non nul; si enfin la fonction $f(x, y)$ est continue sur G, alors on a la formule de changement des variables* (5).

Nous glisserons sur la démonstration du théorème qui est assez compliquée *).

L'intégration par changement de variables d'une intégrale, qu'elle soit définie ou double, est une méthode très efficace qui permet de ramener les intégrales à une forme plus commode pour les calculs.

Exemple 1. Calculer l'intégrale $\int\int_G (2x - y)\,dx\,dy$, où G est le parallélogramme limité par les droites $x + y = 1$, $x + y = 2$, $2x - y = 1$, $2x - y = 3$ (fig. 172, *a*).

Solution. Le calcul direct de cette intégrale est assez laborieux, puisque pour la ramener à une intégrale itérée (par rapport à y et ensuite à x) il faut diviser le domaine G en trois domaines (limités par les pointillés sur la figure 172) et calculer ensuite trois intégrales. Ce problème se simplifie considérablement si l'on envisage le changement de variables

$$x + y = u, \quad 2x - y = v. \tag{6}$$

Les droites $x + y = 1$ et $x + y = 2$ se transforment en $u = 1$ et $u = 2$, les droites $2x - y = 1$ et $2x - y = 3$ en $v = 1$ et

*) Signalons également que la formule (5) est valable dans un cas plus général; en particulier, le jacobien (4) peut s'annuler en un nombre fini de points ou sur des courbes.

$v = 3$ (fig. 172, *b*). Le parallélogramme G se transforme en un rectangle G^* qui constitue un domaine d'intégration plus sim-

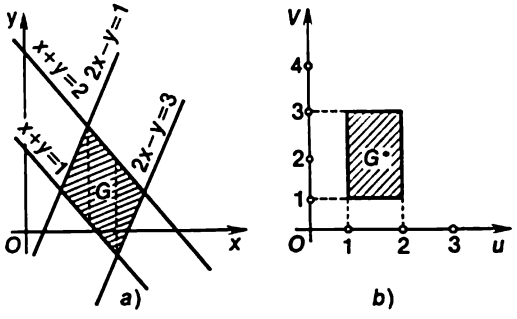


Fig. 172

ple. Reste à déterminer le jacobien. On a $x = (u + v)/3$, $y = (2u - v)/3$. Donc

$$\frac{D(x, y)}{D(u, v)} = \begin{vmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{vmatrix} = -\frac{1}{9} - \frac{2}{9} = -\frac{1}{3}.$$

La formule (5) nous donne finalement

$$\iint_G (2x - y)\,dx\,dy = \iint_{G^*} \frac{1}{3} v\,du\,dv = \frac{1}{3}\int_1^2 du \int_1^3 v\,dv =$$

$$= \frac{1}{3}\int_1^2 \left[\frac{v^2}{2}\right]_1^3 du = \frac{1}{3}\left(\frac{9}{2} - \frac{1}{2}\right)\int_1^2 du = \frac{1}{3}\cdot 4 \cdot 1 = \frac{4}{3}.$$

Remarque. Si l'intégrant ou l'équation de la frontière du domaine d'intégration contient la somme $x^2 + y^2$, il est parfois payant de passer aux coordonnées polaires, car si $x = \rho\cos\varphi$ et $y = \rho\sin\varphi$, cette somme prend la forme simple $(\rho\cos\varphi)^2 + (\rho\sin\varphi)^2 = \rho^2$.

Exemple 2. Calculer l'intégrale $\iint_G e^{x^2+y^2}\,dx\,dy$, où G est la portion de cercle contenue dans le premier quadrant (fig. 173).

Solution. Passons aux coordonnées polaires avec les formules $x = \rho\cos\varphi$, $y = \rho\sin\varphi$. Alors $x^2 + y^2 = \rho^2$ et

$$\frac{D(x, y)}{D(\rho, \varphi)} = \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{dx}{\partial \varphi} \\ \frac{dy}{\partial \rho} & \frac{\partial y}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \cos\varphi & -\rho\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{vmatrix} = \rho[\cos^2\varphi + \sin^2\varphi] = \rho.$$

On voit que ρ varie entre 0 et 1 et φ entre 0 et $\pi/2$. Autrement dit, le domaine G se transforme en le rectangle $\{(\rho; \varphi) \mid 0 \leqslant \rho \leqslant 1; 0 \leqslant \varphi \leqslant \pi/2\}$ *) (fig. 173). Donc la formule (5) nous donne

$$\iint_G e^{x^2+y^2}\,dx\,dy = \int_0^{\pi/2} d\varphi \int_0^1 e^{\rho^2}\rho\,d\rho = \frac{1}{2}(e-1)\int_0^{\pi/2} d\varphi =$$

$$= \frac{1}{2}(e-1)\frac{\pi}{2} = \frac{\pi}{4}(e-1).$$

La construction détaillée du domaine G^* est superflue lorsqu'on change de variables. On se contente généralement de déterminer

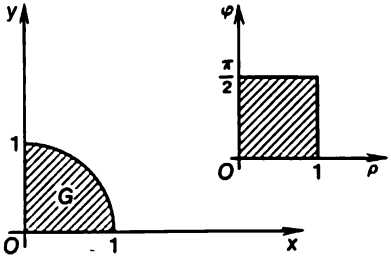


Fig. 173

les limites de variation des nouvelles variables en s'inspirant de la forme du domaine G sur le plan Oxy. C'est d'ailleurs ce qu'on a fait au début de cet exemple.

§ 4. Quelques applications géométriques et physiques des intégrales doubles

1. Calcul du volume. On sait que le volume v d'un cylindre curviligne limité par une surface d'équation $z = f(x, y)$, par le plan Oxy et par un cylindre de génératrice parallèle à Oz et de directrice le contour de G, est donné par la formule

$$v = \iint_G f(x, y)\,dx\,dy.$$

Exemple 1. Calculer le volume limité par $x = 0$, $y = 0$, $z = 0$ et $x + y + z = 1$ (fig. 174).

Solution. On a

$$v = \iint_G (1 - x - y)\,dx\,dy,$$

*) Le jacobien est nul pour $\rho = 0$, mais la formule de changement des variables reste en vigueur.

où G est un triangle limité par les droites $x = 0$, $y = 0$ et $x + y = 1$. Les limites d'intégration sont 0 et 1 pour x, 0 et $1 - x$ pour y. Donc

$$v = \int_0^1 dx \int_0^{1-x} (1 - x - y)\, dy = \int_0^1 \left[(1-x)\, y - \frac{y^2}{2}\right]_0^{1-x} dx =$$

$$= \frac{1}{2} \int_0^1 (1-x)^2\, dx = \frac{1}{6}\,.$$

2. Calcul de l'aire. Comme déjà établi (cf. remarque du § 1), l'aire s d'un domaine G peut être calculée à l'aide de l'intégrale double

$$s = \iint_G dx\, dy.$$

Cette formule est plus universelle que la formule définissant l'aire d'un trapèze curviligne par une intégrale définie, car elle est

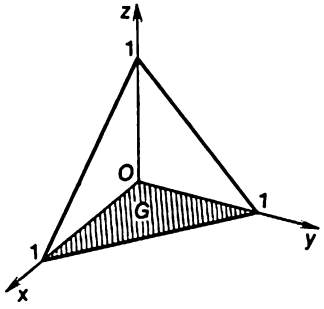


Fig. 174

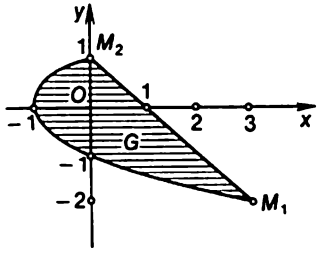


Fig. 175

valable non seulement pour les trapèzes curvilignes, mais aussi pour les figures arbitrairement disposées par rapport aux axes de coordonnées.

Exemple 2. Calculer l'aire du domaine G limité par les courbes $y^2 = x + 1$ et $x + y = 1$ (fig. 175).

Solution. Le domaine G est limité par la parabole $y^2 = x + 1$ et par la droite $y = -x + 1$. La droite coupe la parabole en M_1 (3; −2) et M_2 (0; 1). Donc l'aire cherchée est

$$s = \iint_G dx\, dy = \int_{-2}^{1} dy \int_{y^2-1}^{1-y} dx = \int_{-2}^{1} (2 - y - y^2)\, dy = \frac{9}{2}\,.$$

La détermination des limites d'intégration de l'intégrale itérée est une phase délicate dans le calcul des intégrales doubles. Si

dans cet exemple on avait interverti l'ordre de l'intégration itérée (c'est-à-dire si on avait intégré d'abord par rapport à y et ensuite par rapport à x), on aurait dû partager le domaine G en deux parties (à l'aide de l'axe Oy), puisqu'il est limité par une courbe représentée par des équations différentes sur les intervalles $-1 \leqslant x \leqslant 0$ et $0 \leqslant x \leqslant 3$. On aurait certes obtenu le même résultat mais au prix de calculs bien plus laborieux.

Il est donc utile de se rappeler la règle suivante : si toutes les droites parallèles à l'axe Ox pénètrent dans le domaine d'intégration G par une courbe définie par une seule équation (par la parabole dans le cas de l'exemple 2) et en sortent par une courbe définie par une seule équation (par la droite dans le cas de l'exemple 2), l'intégrale intérieure doit être prise par rapport à x et l'intégrale extérieure, par rapport à y : dans ce cas il n'est pas nécessaire de partager le domaine d'intégration ; de façon analogue, si toutes les droites parallèles à l'axe Oy pénètrent dans le domaine d'intégration G par une courbe définie par une seule équation et en sortent par une courbe définie par une seule équation, l'intégrale intérieure doit être prise par rapport à y et l'intégrale extérieure, par rapport à x.

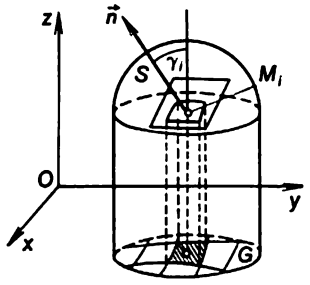


Fig. 176

3. Calcul de l'aire d'une surface. Les intégrales doubles permettent aussi de calculer les aires des surfaces courbes.

Soient S une surface d'équation $z = f(x, y)$, G sa projetée sur le plan Oxy (fig. 176). Supposons que la fonction $f(x, y)$ est continue et admet des dérivées partielles $f'_x(x, y)$ et $f'_y(x, y)$ continues sur G. Pour déterminer l'aire de la surface S, partageons le domaine G arbitrairement en n parties G_i sans points intérieurs communs, d'aires Δs_i $(i = 1, 2, \ldots, n)$, et désignons par S_i l'élément de surface S dont le projeté sur le plan Oxy est l'élément G_i. La surface S est donc partagée en n parties.

A chaque point $(\xi_i ; \eta_i)$ choisi arbitrairement dans G_i correspond un point $M_i [\xi_i ; \eta_i ; f(\xi_i, \eta_i)]$ de S. Menons par M_i un plan tangent à S. L'équation de ce plan est

$$f'_x(\xi_i, \eta_i)(x - \xi_i) + f'_y(\xi_i, \eta_i)(y - \eta_i) - (z - z_i) = 0,$$

où x, y, z sont les coordonnées du point courant du plan ; ξ_i, η_i, $z_i = f(\xi_i, \eta_i)$ les coordonnées du point de contact (cf. chap. XII, § 4, n° 2). On rappelle que le vecteur normal $\vec{n}$ du plan tangent

a pour composantes :

$$\{-f'_x(\xi_i, \eta_i);\ -f'_y(\xi_i, \eta_i);\ +1\}.$$

Considérons la partie du plan tangent dont la projetée sur le plan Oxy est l'élément G_i. Désignons-la par σ_i et son aire par $\Delta\sigma_i$. On peut considérer que l'aire $\Delta\sigma_i$ est approximativement égale à l'aire de S_i et la somme de toutes les aires

$$\sum_{i=1}^{n} \Delta\sigma_i,$$

approximativement égale à l'aire de S.

La valeur exacte de l'aire s de S sera par définition posée égale à la limite de la somme

$$s = \lim_{\lambda \to 0} \sum_{i=1}^{n} \Delta\sigma_i^{*}), \tag{1}$$

où λ est le plus grand des diamètres des éléments G_i. Montrons que cette limite existe et qu'elle est égale à l'intégrale double

$$s = \iint_G \sqrt{1 + f'^2_x(x, y) + f'^2_y(x, y)}\, dx\, dy. \tag{2}$$

Désignons par γ_i l'angle $(\widehat{\overrightarrow{Oz}, \overrightarrow{n}})$. Cet angle est égal à celui du plan tangent en M_i à S et du plan Oxy. L'élément G_i étant le projeté de σ_i sur le plan Oxy, les aires de ces éléments sont reliées par la relation

$$\Delta\sigma_i = \frac{\Delta s_i}{\cos\gamma_i}.$$

En effet, cette formule est, on le sait, valable pour les triangles. Elle le sera de toute évidence et pour les polygones plans, car ces derniers peuvent être partagés en plusieurs triangles, et pour toute figure plane d'aire $\Delta\sigma$, limitée par une courbe, car son aire peut être traitée comme la limite des aires de polygones inscrits.

D'autre part, on sait du cours de géométrie analytique que

$$\cos\gamma_i = \frac{1}{\sqrt{1 + f'^2_x(\xi_i, \eta_i) + f'^2_y(\xi_i, \eta_i)}}.$$

Donc

$$\Delta\sigma_i = \sqrt{1 + f'^2_x(\xi_i, \eta_i) + f'^2_y(\xi_i, \eta_i)}\, \Delta s_i.$$

---

*) La limite de la somme $\sum_{i=1}^{n} \Delta\sigma_i$ lorsque $\lambda \to 0$ se définit comme celle de la somme intégrale d'une intégrale double (cf. note de la page 344). Dans la suite toutes les définitions identiques seront omises.

En portant $\Delta\sigma_i$ dans la somme (1), on trouve

$$s = \lim_{\lambda \to 0} \sum_{i=1}^{n} \sqrt{1 + f'^2_x(\xi_i, \eta_i) + f'^2_y(\xi_i, \eta_i)}\, \Delta s_i.$$

La somme du second membre est une somme intégrale pour la fonction $\sqrt{1 + f'^2_x(x, y) + f'^2_y(x, y)}$. Cette fonction étant par hypothèse continue sur le domaine G, la limite de cette somme lorsque $\lambda \to 0$ existe et est égale à l'intégrale double (2). C.Q.F.D.

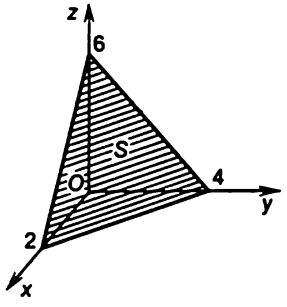


Fig. 177

La formule (2) permet de calculer l'aire des surfaces définies par une équation $z = f(x, y)$.

Exemple 3. Calculer l'aire de la région du plan $6x + 3y + 2z = 12$ comprise dans le premier octant (fig. 177).

Solution. Puisque la fonction $z = 6 - 3x - \frac{3}{2}y$ et le domaine G, projeté de la région envisagée sur le plan Oxy, satisfont aux conditions énoncées ci-dessus, l'aire peut être calculée à l'aide de la formule (2). On a

$$f'_x(x, y) = -3,\ f'_y(x, y) = -3/2;$$
$$\sqrt{1 + f'^2_x(x, y) + f'^2_y(x, y)} = \sqrt{1 + 9 + 9/4} = 7/2.$$

Le domaine G est un triangle limité par les axes Ox et Oy et par la droite $6x + 3y = 12$, projeté du plan donné sur Oxy. Les limites d'intégration sont 0 et 2 pour x et 0 et $4 - 2x$ pour y. Donc

$$s = \int_0^2 dx \int_0^{4-2x} \frac{7}{2}\, dy = \frac{7}{2} \int_0^2 [y]_0^{4-2x}\, dx = \frac{7}{2} \int_0^2 (4 - 2x)\, dx =$$
$$= \frac{7}{2} [4x - x^2]_0^2 = \frac{7}{2} \cdot 4 = 14.$$

4. Calcul de la masse d'une plaque. Considérons une plaque dans le plan Oxy, c'est-à-dire un domaine G dont la masse m est distribuée avec une densité $\rho(x, y)$. Calculons la masse m de cette plaque à l'aide de la densité $\rho(x, y)$ en admettant que $\rho(x, y)$ est une fonction continue. Partageons G arbitrairement en n éléments G_i $(i = 1, 2, \ldots, n)$ et désignons par m_i la masse de G_i. Prenons un point arbitraire $(\xi_i; \eta_i)$ dans chaque élément G_i. La masse m_i peut être supposée approximativement égale à $\rho(\xi_i, \eta_i)\, \Delta s_i$, où Δs_i est l'aire de G_i, et la masse totale m de la

plaque, approximativement égale à la somme

$$m = \sum_{i=1}^{n} m_i \approx \sum_{i=1}^{n} \rho(\xi_i, \eta_i)\,\Delta s_i,$$

qui est une somme intégrale pour la fonction continue $\rho(x, y)$ dans le domaine G. En faisant tendre λ vers 0, on trouve de toute évidence la valeur exacte de la masse de la plaque, valeur qui est égale à l'intégrale double de la fonction $\rho(x, y)$ étendue au domaine G, c'est-à-dire à

$$m = \iint_G \rho(x, y)\,dx\,dy. \quad (3)$$

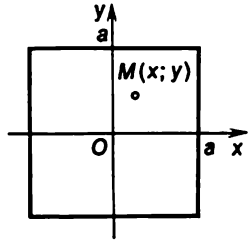


Fig. 178

Exemple 4. Calculer la masse d'une plaque carrée de côté $2a$ dont la densité $\rho(x, y)$ en chaque point $M(x; y)$ est proportionnelle au carré de la distance de M au point d'intersection des diagonales, le coefficient de proportionnalité étant k.

Solution. Choisissons le système de coordonnées comme indiqué sur la figure 178. Cherchons ensuite la fonction $\rho(x, y)$ à partir des conditions du problème. Soit $M(x; y)$ un point arbitraire de la plaque. Le carré de la distance du point M au point d'intersection des diagonales est égal à $x^2 + y^2$. Donc la densité en M est

$$\rho(M) = \rho(x, y) = k(x^2 + y^2).$$

D'après la formule (3)

$$m = \iint_G k(x^2 + y^2)\,dx\,dy.$$

L'intégrant étant pair par rapport à x et y, et le domaine d'intégration, symétrique par rapport aux axes de coordonnées, on peut étendre l'intégration à la portion de domaine G comprise dans le premier quadrant, c'est-à-dire

$$m = 4k \int_0^a dx \int_0^a (x^2 + y^2)\,dy = 4k \int_0^a \left[x^2 y + \frac{y^3}{3}\right]_0^a dx =$$

$$= 4k \int_0^a \left(ax^2 + \frac{a^3}{3}\right) dx = 4k \left[\frac{ax^3}{3} + \frac{a^3 x}{3}\right]_0^a = 4k \frac{2a^4}{3} = \frac{8}{3} ka^4.$$

5. Centre de gravité d'une plaque. Trouvons les coordonnées du centre de gravité d'une plaque occupant un domaine G du

plan Oxy. Soit $\rho(x, y)$, où $\rho(x, y)$ est une fonction continue, la densité de cette plaque en $M(x; y)$. Partageons le domaine G en éléments G_i $(i = 1, 2, \ldots, n)$, choisissons un point $(\xi_i; \eta_i)$ dans chacun de ces éléments et admettons que la masse m_i de G_i est approximativement égale à $\rho(\xi_i, \eta_i)\,\Delta s_i$, où Δs_i est l'aire de G_i. Si l'on admet que chacune de ces masses est concentrée en un point, plus exactement au point $(\xi_i; \eta_i)$, les coordonnées x_c et y_c du barycentre d'un tel système de points matériels seront définies par les formules

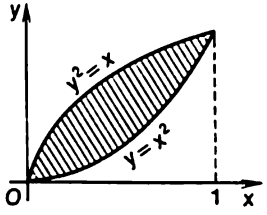


Fig. 179

$$x_c = \frac{\sum_{i=1}^{n} \xi_i \rho(\xi_i, \eta_i)\,\Delta s_i}{\sum_{i=1}^{n} \rho(\xi_i, \eta_i)\,\Delta s_i}; \qquad y_c = \frac{\sum_{i=1}^{n} \eta_i \rho(\xi_i, \eta_i)\,\Delta s_i}{\sum_{i=1}^{n} \rho(\xi_i, \eta_i)\,\Delta s_i}, \tag{4}$$

formules qui donnent approximativement les coordonnées du centre de gravité de la plaque. Pour obtenir les valeurs exactes de ces coordonnées, il faut faire tendre λ vers 0 dans (4). On obtient alors

$$x_c = \frac{\iint\limits_G x\rho(x, y)\,dx\,dy}{m}; \qquad y_c = \frac{\iint\limits_G y\rho(x, y)\,dx\,dy}{m}, \tag{5}$$

où $m = \iint\limits_G \rho(x, y)\,dx\,dy$ est la masse de la plaque.

Si la plaque est homogène, c'est-à-dire que $\rho = \text{const}$, les formules du centre de gravité se simplifient :

$$x_c = \frac{\iint\limits_G x\,dx\,dy}{\iint\limits_G dx\,dy}; \qquad y_c = \frac{\iint\limits_G y\,dx\,dy}{\iint\limits_G dx\,dy}. \tag{6}$$

Les quantités $M_y = \iint\limits_G x\rho(x, y)\,dx\,dy$ et $M_x = \iint\limits_G y\rho(x, y)\,dx\,dy$ s'appellent *moments statiques* de la plaque par rapport aux axes Oy et Ox.

Le calcul des coordonnées du centre de gravité se ramène donc à celui de trois intégrables doubles.

Exemple 5. Trouver le centre de gravité d'une plaque homogène limitée par les deux paraboles $y^2 = x$ et $x^2 = y$ (fig. 179).

S o l u t i o n. Déterminons les coordonnées du centre de gravité de cette plaque à l'aide des formules (6). Calculons tout d'abord la masse de cette plaque:

$$m = \int_G \int dx\, dy = \int_0^1 dx \int_{x^2}^{\sqrt{x}} dy = \frac{1}{3}.$$

Calculons ensuite ses moments statiques par rapport aux axes de coordonnées:

$$M_y = \int_G \int x\, dx\, dy = \int_0^1 x\, dx \int_{x^2}^{\sqrt{x}} dy = \frac{3}{20};$$

$$M_x = \int_G \int y\, dx\, dy = \int_0^1 dx \int_{x^2}^{\sqrt{x}} y\, dy = \frac{3}{20}.$$

Les formules (6) nous donnent maintenant

$$x_c = \frac{M_y}{m} = \frac{3}{20} : \frac{1}{3} = \frac{9}{20}; \qquad y_c = \frac{M_x}{m} = \frac{3}{20} : \frac{1}{3} = \frac{9}{20}.$$

Ainsi $x_c = y_c = \frac{9}{20}$.

6. Moment d'inertie d'une plaque. On sait que le moment d'inertie d'un point matériel par rapport à un axe est égal au produit de la masse de ce point par le carré de sa distance à l'axe, et le moment d'inertie d'un système de points matériels, à la somme des moments d'inertie de ces points.

Supposons qu'un domaine G du plan Oxy est occupé par une plaque de densité continue $\rho(x, y)$. Partageons le domaine G en éléments G_i d'aires Δs_i ($i = 1, 2, \ldots, n$) et choisissons un point $(\xi_i; \eta_i)$ dans chaque élément G_i. Nous remplaçons ainsi la plaque par un système de points matériels de masses $m_i = \rho(\xi_i, \eta_i)\,\Delta s_i$ et de coordonnées $(\xi_i; \eta_i)$. Le moment d'inertie d'un tel système de points pesants par rapport par exemple à Oy est égal à $\sum_{i=1}^{n} \xi_i^2 \rho(\xi_i, \eta_i)\,\Delta s_i$. Prenons cette expression pour valeur approchée du moment d'inertie de la plaque. Or, cette expression est une somme intégrale pour la fonction continue $x^2\rho(x, y)$. En faisant tendre λ vers 0, on obtient pour le moment d'inertie de la plaque par rapport à Oy la formule suivante:

$$I_y = \int_G \int x^2 \rho(x, y)\, dx\, dy.$$

De façon analogue, le moment d'inertie par rapport à l'axe Ox est égal à

$$I_x = \int_G \int y^2 \rho(x, y)\, dx\, dy.$$

Trouvons le moment d'inertie I_0 de la plaque par rapport à l'origine des coordonnées. Puisque le moment d'inertie d'un point matériel de masse m par rapport à l'origine des coordonnées est égal à $m(x^2 + y^2)$, on obtient en reprenant les raisonnements précédents

$$I_0 = \int_G \int (x^2 + y^2) \rho(x, y)\, dx dy, \tag{7}$$

c'est-à-dire que

$$I_0 = I_x + I_y.$$

Exemple 6. Trouver le moment d'inertie d'un disque de rayon R de densité constante $\rho(x, y) = 1$ par rapport à l'origine des coordonnées.

Solution. La formule (7) nous donne

$$I_0 = \int_G \int (x^2 + y^2)\, dx\, dy.$$

Passons aux coordonnées polaires. L'équation du contour du disque étant $\rho = R$ en coordonnées polaires, on a

$$I_0 = \int_0^{2\pi} d\varphi \int_0^R \rho^2 \rho\, d\rho = \int_0^{2\pi} \left[\frac{\rho^4}{4}\right]_0^R d\varphi = \frac{1}{4}[R^4\varphi]_0^{2\pi} = \frac{\pi R^4}{2}.$$

§ 5. Intégrales curvilignes

Généralisons la notion d'intégrale définie au cas où le domaine d'intégration est un segment de courbe plane.

Les intégrales de cette espèce sont dites *curvilignes*. Elles interviennent largement dans les diverses branches des mathématiques.

On distingue deux types d'intégrales curvilignes: les intégrales de première et de seconde espèce.

1. Définition de l'intégrale curviligne de première espèce. Considérons dans le plan Oxy une courbe AB différentiable ou différentiable par morceaux *) et soit $z = f(x, y)$ une fonction définie et bornée sur la courbe AB.

*) Une courbe définie par les équations $x = \varphi(t)$, $y = \psi(t)$, $\alpha \leqslant t \leqslant \beta$, est dite *différentiable* si les fonctions $\varphi(t)$ et $\psi(t)$ sont continues et possèdent des dérivées premières continues non toutes deux nulles (ce qui exprime que cette courbe admet une tangente en chacun de ses points). Une courbe continue composée d'un nombre fini de portions différentiables est dite *différentiable par morceaux*.

Partageons arbitrairement la courbe AB en n parties par les points $A = M_0, M_1, M_2, \ldots, M_{i-1}, M_i, \ldots, M_{n-1}, M_n = B$, chosissons un point arbitraire M_i^* dans chaque arc $M_{i-1}M_i$ (fig. 180) et composons la somme

$$\sum_{i=1}^{n} f(M_i^*)\,\Delta l_i, \tag{1}$$

où Δl_i est la longueur de l'arc $M_{i-1}M_i$. La somme (1) s'appelle *somme intégrale* pour la fonction $z = f(x, y) = f(M)$ le long de la courbe AB. Désignons par λ la plus grande des longueurs des arcs $M_{i-1}M_i$ ($\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta l_i\}$).

Définition. La limite I, si elle existe, de la somme intégrale (1) lorsque $\lambda \to 0$ s'appelle *intégrale curviligne de première espèce* de la fonction $f(x, y)$ le long de la courbe AB et se représente par l'une des notations suivantes

$$I = \int\limits_{\widehat{AB}} f(M)\,dl = \int\limits_{\widehat{AB}} f(x, y)\,dl.$$

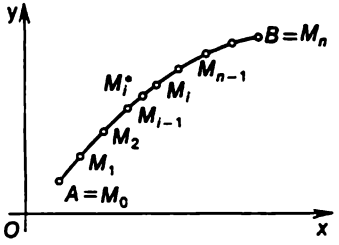


Fig. 180

La fonction $f(x, y)$ est dite *intégrable* le long de la courbe AB, la courbe AB est le *contour d'intégration*; les points A et B, *bornes d'intégration*.

L'intégrale curviligne de première espèce se ramène sans peine à l'intégrale définie. En effet, en paramétrant la courbe AB à l'aide d'un paramètre l égal à la longueur d'arc mesurée à partir de A, on obtient la représentation paramétrique de la courbe: $x = x(l)$, $y = y(l)$ $(0 \leqslant l \leqslant L)$. Désignons par l_i^* la valeur du paramètre l correspondant au point M_i^*, par l_i, la valeur correspondant à M_i et mettons la somme intégrale (1) sous la forme

$$\sum_{i=1}^{n} f[x(l_i^*), y(l_i^*)]\,\Delta l_i, \tag{2}$$

où $\Delta l_i = l_i - l_{i-1}$ et $l_{i-1} \leqslant l_i^* \leqslant l_i$. La somme (2) est somme intégrale pour la fonction $f[x(l), y(l)]$ sur l'intervalle $[0, L]$. Les sommes intégrales (1) et (2) étant égales, il en est de même des intégrales qui leur correspondent, c'est-à-dire

$$\int\limits_{\widehat{AB}} f(x, y)\,dl = \int\limits_{0}^{L} f[x(l), y(l)]\,dl. \tag{3}$$

Remarquons que la formule (3) ne fait pas qu'exprimer une intégrale curviligne en fonction d'une intégrale définie, elle

prouve l'existence de l'intégrale curviligne d'une fonction $f(x, y)$ continue le long de la courbe AB *).

Comme déjà indiqué, l'intégrale curviligne de première espèce se ramène directement à l'intégrale définie, mais entre ces deux notions nous avons la distinction suivante. Dans la somme intégrale (1) les quantités Δl_i sont nécessairement >0, indépendamment des points pris pour origine et extrémité de la courbe AB, c'est-à-dire que

$$\int_{\widehat{AB}} f(x, y)\, dl = \int_{\widehat{BA}} f(x, y)\, dl,$$

tandis que l'intégrale définie $\int_a^b f(x)\, dx$ change de signe si l'on intervertit les bornes d'intégration. Pour le reste ces intégrales sont douées des mêmes propriétés. Ceci résulte directement de la formule (3).

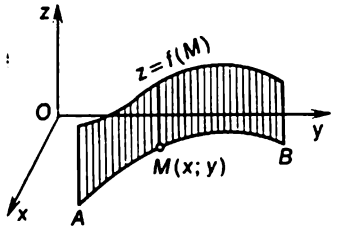


Fig. 181

L'intégrale curviligne de première espèce admet aussi une interprétation géométrique. Si l'intégrale définie $\int_a^b f(x)\, dx$ où $f(x) \geqslant 0$ est l'aire d'un trapèze curviligne, l'intégrale curviligne $\int_{\widehat{AB}} f(M)\, dl$ où $f(M) \geqslant 0$ est numériquement égale à l'aire de la portion de surface cylindrique engendrée par les perpendiculaires au plan Oxy, de longueur variable $f(M)$, s'appuyant sur la courbe AB (fig. 181).

En particulier, si AB n'est pas une courbe, mais un segment $[a, b]$ de l'axe Ox, alors $f(x, y) = f(x)$, $\Delta l_i = \Delta x_i$ et l'intégrale curviligne est une ordinaire intégrale définie.

Si, enfin, on pose $f(M) \equiv 1$, on obtient l'intégrale curviligne $\int_{\widehat{AB}} dl$ qui est égale à la longueur de l'arc $\widehat{AB}$.

Ainsi l'intégrale curviligne de première espèce permet de calculer l'aire des surfaces cylindriques et les longueurs des arcs. Cette intégrale possède par ailleurs un vaste champ d'application en physique. Elle peut servir au même titre que l'intégrale double à calculer la masse d'une courbe matérielle à l'aide de sa densité, les moments d'inertie par rapport aux axes de coordonnées, les coordonnées du centre de gravité d'une courbe matérielle, etc.

*) Dire que $f(x, y) = f(M)$ est continue le long de la courbe AB revient à dire que $\lim_{M \to M_0} f(M) = f(M_0)$ en tout point M_0 de AB où $M \in AB$.

2. Calcul des intégrales curvilignes de première espèce. Le calcul des intégrales curvilignes de première espèce se ramène à celui d'intégrales définies.

Soient AB une courbe d'équations paramétriques $x = \varphi(t)$, $y = \psi(t)$ $(\alpha \leqslant t \leqslant \beta)$, où $\varphi(t)$ et $\psi(t)$ sont des fonctions continues avec leurs dérivées premières; $f(x, y)$ une fonction continue sur AB. Nous admettrons pour fixer les idées qu'aux points A et B correspondent respectivement les valeurs α et β de t. Pour tout point $M(\varphi(t); \psi(t))$ de la courbe AB, la longueur l de l'arc $\overset{\frown}{AM}$ peut être traitée comme une fonction du paramètre t et calculée (chap. VIII, § 10, n° 3) par la formule

$$l = l(t) = \int_{\alpha}^{t} \sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2}\, dt,$$

d'où il s'ensuit, en vertu de la règle de dérivation d'une intégrale par rapport à la borne supérieure,

$$dl = \sqrt{[(\varphi'(t)]^2 + [\psi'(t)]^2}\, dt. \tag{4}$$

En tenant compte de (4), on peut mettre (3) sous la forme

$$\int_{\overset{\frown}{AB}} f(x, y)\, dl = \int_{0}^{L} f[x(l), y(l)]\, dl =$$

$$= \int_{\alpha}^{\beta} f[\varphi(t), \psi(t)] \sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2}\, dt. \tag{5}$$

Exemple 1. Calculer l'intégrale curviligne $\int_{\overset{\frown}{AB}} y^2\, dl$, où $\overset{\frown}{AB}$ est le quart de cercle $x = a\cos t$, $y = a\sin t$, $0 \leqslant t \leqslant \pi/2$.

Solution. Puisque

$$y^2 = a^2 \sin^2 t,\quad dl = \sqrt{a^2 \sin^2 t + a^2 \cos^2 t}\, dt = a\, dt,$$

la formule (5) nous donne

$$\int_{\overset{\frown}{AB}} y^2\, dl = \int_{0}^{\pi/2} a^2 \sin^2 t \cdot a\, dt = \frac{a^3}{2} \int_{0}^{\pi/2} (1 - \cos 2t)\, dt =$$

$$= \frac{a^3}{2}\left[t - \frac{\sin 2t}{2}\right]_0^{\pi/2} = \frac{a^3 \pi}{4}.$$

En particulier, si la courbe AB est définie par l'équation $y = y(x)$, $a \leqslant x \leqslant b$, où $y(x)$ est une fonction continûment dérivable, en prenant x pour paramètre, on obtient à partir de la

formule (5)

$$\int\limits_{\overset{\frown}{AB}} f(x,\ y)\,dl = \int\limits_{a}^{b} f[x,\ y(x)]\,\sqrt{1+y'^2(x)}\,dx. \qquad (6)$$

Exemple 2. Calculer l'intégrale curviligne $\int\limits_{\overset{\frown}{AB}} y\,dl$, où $\overset{\frown}{AB}$ est l'arc de parabole $y^2 = 2x$ d'extrémités (0; 0) et (2; 2).

Solution. On a

$$y = \sqrt{2x},\ y' = 1/\sqrt{2x},\ dl = \sqrt{1+y'^2}\,dx = \sqrt{1+1/(2x)}\,dx.$$

La formule (6) nous donne

$$\int\limits_{\overset{\frown}{AB}} y\,dl = \int\limits_{0}^{2} \sqrt{2x}\sqrt{1+1/(2x)}\,dx =$$

$$= \int\limits_{0}^{2} \sqrt{2x+1}\,dx = \frac{1}{3}\left[(2x+1)^{3/2}\right]_0^2 = \frac{1}{3}(5\sqrt{5}-1).$$

Remarque. La formule (4) est intéressante en soi. En élevant au carré, on obtient : $(dl)^2 = [\varphi'(t)\,dt]^2 + [\psi'(t)\,dt]^2 = (dx)^2 + (dy)^2$. Cette égalité nous suggère une interprétation géométrique simple de la différentielle d'arc dl. Puisque la différentielle de la fonction $y = y(x)$ est égale à l'accroissement de l'ordonnée de la tangente (chap. V, § 3, n° 1), la différentielle dl (cf. fig. 185) est égale à la longueur du segment de tangente à la courbe AB au point d'abscisse x, d'extrémités les points $(x;\ y)$ et $(x + dx;\ y + dy)$, c'est-à-dire à l'hypoténuse du triangle rectangle de côtés $|\,dx\,|$ et $|\,dy\,|$; l'égalité $(dl)^2 = (dx)^2 + (dy)^2$ est la formule du triangle rectangle (théorème de Pythagore).

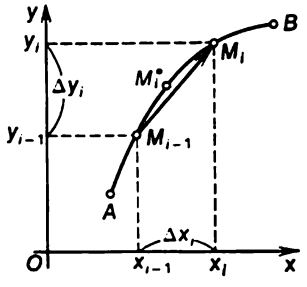


Fig. 182

3. Définition de l'intégrale curviligne de seconde espèce. Soient données sur une courbe AB deux fonctions bornées $P(x, y)$ et $Q(x, y)$. Partageons la courbe AB en n parties par les points $A = M_0, M_1, \ldots, M_{i-1}, M_i, \ldots, M_n = B$. Désignons par Δx_i et Δy_i les mesures algébriques des projetés du vecteur $\overrightarrow{M_{i-1}M_i}$ sur les axes de coordonnées (fig. 182), prenons un point arbitraire M_i^* dans chaque arc $\overset{\frown}{M_{i-1}M_i}$ et composons la somme intégrale

pour la fonction $P(x, y)$ (resp. $Q(x, y)$) :

$$\sum_{i=1}^{n} P(M_i^*)\,\Delta x_i \left(\text{resp.} \sum_{i=1}^{n} Q(M_i^*)\Delta y_i\right). \tag{7}$$

Définition. La limite I, si elle existe, de la somme intégrale (7) lorsque $\lambda \to 0$ ($\lambda = \max\limits_{1 \leqslant i \leqslant n} \{\Delta l_i\}$; Δl_i étant la longueur de l'arc $\overset{\frown}{M_{i-1}M_i}$), s'appelle *intégrale curviligne de seconde espèce* de la fonction $P(x, y)$ (resp. $Q(x, y)$) le long de la courbe AB et se note

$$\int_{\overset{\frown}{AB}} P(x, y)\,dx \quad \left(\text{resp.} \int_{\overset{\frown}{AB}} Q(x, y)\,dy\right).$$

La somme

$$\int_{\overset{\frown}{AB}} P(x, y)\,dx + \int_{\overset{\frown}{AB}} Q(x, y)\,dy$$

s'appelle *intégrale curviligne générale de seconde espèce* et se représente par

$$\int_{\overset{\frown}{AB}} P(x, y)\,dx + Q(x, y)\,dy \text{ *)}.$$

Les intégrales curvilignes de seconde espèce se ramènent facilement aux intégrales définies.

En effet, supposons que la courbe AB est représentée par les équations paramétriques $x = \varphi(t)$, $y = \psi(t)$, $\alpha \leqslant t \leqslant \beta$, où $\varphi(t)$ et $\psi(t)$ sont continues avec leurs dérivées premières, que $\varphi'^2(t) + \psi'^2(t) \neq 0$ et qu'aux points A et B correspondent les valeurs α et β de t. Soient $P(x, y)$ et $Q(x, y)$ des fonctions continues le long de la courbe AB. On a alors les formules suivantes

$$\begin{aligned}
&\int_{\overset{\frown}{AB}} P(x, y)\,dx = \int_{\alpha}^{\beta} P[\varphi(t), \psi(t)]\,\varphi'(t)\,dt;\\
&\int_{\overset{\frown}{AB}} Q(x, y)\,dy = \int_{\alpha}^{\beta} Q[\varphi(t), \psi(t)]\,\psi'(t)\,dt;\\
&\int_{\overset{\frown}{AB}} P(x, y)\,dx + Q(x, y)\,dy =\\
&= \int_{\alpha}^{\beta} \{P[\varphi(t), \psi(t)]\,\psi'(t) + Q[\varphi(t), \psi(t)]\,\psi'(t)\}\,dt,
\end{aligned} \tag{8}$$

*) On omettra parfois les arguments des fonctions $P(x, y)$ et $Q(x, y)$ et on écrira simplement P et Q.

qui ramènent les intégrales curvilignes à des intégrales définies.

Prouvons la première des formules (8):

$$\int_{\widehat{AB}} P(x, y)\,dx = \int_{\alpha}^{\beta} P[\varphi(t), \psi(t)]\,\varphi'(t)\,dt, \tag{9}$$

la seconde se prouvant de façon analogue et la troisième étant la somme des deux premières.

Supposons qu'aux points M_i de subdivision de la courbe AB correspondent les valeurs t_i de t, aux points M_i^*, les valeurs t_i^*, c'est-à-dire que les coordonnées du point M_i sont $[\varphi(t_i); \psi(t_i)]$, celles du point M_i^*, $[\varphi(t_i^*); \psi(t_i^*)]$, $i = 1, 2, \ldots, n$. La fonction $P(x, y)$ est une fonction composée de t: $P[\varphi(t), \psi(t)]$. Les fonctions $x = \varphi(t)$ et $y = \psi(t)$ étant continues sur l'intervalle $[\alpha, \beta]$ et la fonction $P(x, y)$ continue sur AB, le théorème de continuité d'une fonction composée nous dit que la fonction $P[\varphi(t), \psi(t)]$ est continue sur l'intervalle $[\alpha, \beta]$.

Formons la somme intégrale (7) pour la fonction $P(x, y)$:

$$\sigma = \sum_{i=1}^{n} P(M_i^*)\,\Delta x_i = \sum_{i=1}^{n} P[\varphi(t_i^*), \psi(t_i^*)]\,\Delta x_i.$$

Comme $\Delta x_i = \varphi(t_i) - \varphi(t_{i-1})$, la formule de Newton-Leibniz nous donne

$$\Delta x_i = \varphi(t_i) - \varphi(t_{i-1}) = \int_{t_{i-1}}^{t_i} \varphi'(t)\,dt.$$

Donc

$$\sigma = \sum_{i=1}^{n} P[\varphi(t_i^*), \psi(t_i^*)] \int_{t_{i-1}}^{t_i} \varphi'(t)\,dt =$$

$$= \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} P[\varphi(t_i^*), \psi(t_i^*)]\,\varphi'(t)\,dt.$$

D'autre part, la fonction $P[\varphi(t), \psi(t)]\,\varphi'(t)$ étant continue sur $[\alpha, \beta]$, son intégrale définie du second membre de (9) existe. Représentons cette intégrale par la somme d'intégrales sur les intervalles $[t_{i-1}, t_i]$:

$$I = \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} P[\varphi(t), \psi(t)]\,\varphi'(t)\,dt.$$

Estimons la différence

$$\sigma - I = \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \{P[\varphi(t_i^*), \psi(t_i^*)] - P[\varphi(t), \psi(t)]\} \varphi'(t)\, dt. \quad (10)$$

Etant continue sur $[\alpha, \beta]$, la fonction $P[\varphi(t), \psi(t)]$ y est uniformément continue d'après le théorème de Cantor. Or, cela exprime que pour tout $\varepsilon > 0$ on peut exhiber un $\delta > 0$ tel que pour $\mu = \max_{1 \leqslant i \leqslant n} \{\Delta t_i\} < \delta$ l'on ait

$$|P[\varphi(t_i^*), \psi(t_i^*)] - P[\varphi(t), \psi(t)]| < \varepsilon. \quad (11)$$

La fonction $\varphi'(t)$ est bornée sur $[\alpha, \beta]$, puisqu'elle y est continue, c'est-à-dire qu'il existe un nombre k tel que

$$|\varphi'(t)| \leqslant k. \quad (12)$$

En se servant de (11) et (12), on trouve la majoration suivante pour la différence (10):

$$|\sigma - I| \leqslant \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} |P[\varphi(t_i^*), \psi(t_i^*)] - P[\varphi(t), \psi(t)]|\, |\varphi'(t)|\, dt <$$

$$< \varepsilon k \sum_{i=1}^{n} (t_i - t_{i-1}) = \varepsilon k (\beta - \alpha).$$

D'où, puisque ε est arbitraire,

$$\lim_{\mu \to 0} \sigma = I. \quad (13)$$

Mais $\lambda = \max_{1 \leqslant i \leqslant n} \{\Delta t_i\}$ tend aussi vers 0 lorsque $\mu = \max_{1 \leqslant i \leqslant n} \{\Delta l_i\}$ tend vers 0, et inversement. En effet, $\Delta l_i = \int_{t_{i-1}}^{t_i} \sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2}\, dt$. La continuité des fonctions $\varphi'(t)$ et $\psi'(t)$ sur $[\alpha, \beta]$ entraîne celle de la fonction $\sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2}$ sur $[\alpha, \beta]$. Alors $m \Delta t_i \leqslant \Delta l_i \leqslant M \Delta t_i$, où m et M sont le minimum et le maximum de la fonction $\sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2}$ sur $[\alpha, \beta]$ et de plus $m > 0$ et $M > 0$ en vertu de la condition $\varphi'^2(t) + \psi'^2(t) \neq 0$. De l'inégalité de gauche il résulte que $\mu \to 0$ lorsque $\lambda \to 0$ et de l'inégalité de droite, que $\lambda \to 0$ lorsque $\mu \to 0$. Donc de (3) il s'ensuit

$$\lim_{\lambda \to 0} \sigma = I,$$

c'est-à-dire que l'intégrale curviligne $\int_{AB} P(x, y)\, dx$ existe et l'on a la formule (9).

Des formules (8) il résulte immédiatement que l'intégrale curviligne de seconde espèce jouit des mêmes propriétés que l'intégrale curviligne de première espèce.

Contrairement à l'intégrale curviligne de première espèce, l'intégrale de seconde espèce change de signe si la courbe AB est parcourue dans des sens différents, c'est-à-dire que

$$\int_{\widehat{AB}} P(x, y)\,dx = -\int_{\widehat{BA}} P(x, y)\,dx,$$

$$\int_{\widehat{AB}} Q(x, y)\,dy = -\int_{\widehat{BA}} Q(x, y)\,dy.$$

En effet, si l'on change le sens de parcours de la courbe, on change les signes des mesures algébriques Δx_i et Δy_i dans les sommes (7), donc le signe de ces sommes et de leurs limites.

Dans le calcul des intégrales curvilignes de seconde espèce il faut donc tenir compte du sens de parcours de la courbe.

Lorsque L est une courbe fermée, c'est-à-dire lorsque B est confondu avec A, on conviendra que le sens *direct* ou *positif* est celui qui laisse le domaine intérieur limité par L à gauche du point courant se déplaçant sur L. Le sens opposé sera appelé sens *rétrograde* ou *négatif*.

L'intégrale curviligne le long d'un contour fermé L parcouru dans le sens direct est souvent désignée par la notation

$$\oint_L P(x, y)\,dx + Q(x, y)\,dy.$$

4. Calcul des intégrales curvilignes de seconde espèce. Les intégrales curvilignes de seconde espèce se calculent par réduction à des intégrales définies à l'aide des formules (8).

En particulier, si une courbe AB est définie par une équation $y = y(x)$, $a \leqslant x \leqslant b$, où $y(x)$ est une fonction continûment dérivable, en prenant x pour paramètre, on déduit des formules (8) que

$$\int_{\widehat{AB}} P(x, y)\,dx = \int_a^b P[x, y(x)]\,dx,$$

$$\int_{\widehat{AB}} Q(x, y)\,dy = \int_a^b Q[x, y(x)]\,y'(x)\,dx, \tag{14}$$

$$\int_{\widehat{AB}} P(x, y)\,dx + Q(x, y)\,dy = \int_a^b \{P[x, y(x)] + Q[x, y(x)]\,y'(x)\}\,dx.$$

On obtient des formules identiques si la courbe AB est définie par une équation de la forme $x = x(y)$.

Exemple 3. Calculer l'intégrale $\int\limits_{\overset{\frown}{B}} x^2\,dx + xy\,dy$, où $\overset{\frown}{AB}$ est le quart de cercle $x = \cos t$, $y = \sin t$, $0 \leqslant t \leqslant \pi/2$, aux points A et B correspondent les valeurs 0 et $\pi/2$ de t.

Solution. On a $x^2 = \cos^2 t$, $dx = -\sin t\,dt$, $xy = \cos t \sin t$, $dy = \cos t\,dt$. La troisième formule (8) nous donne

$$\int\limits_{\overset{\frown}{AB}} x^2\,dx + xy\,dy = \int\limits_0^{\pi/2} (-\cos^2 t \sin t + \cos^2 t \sin t)\,dt = 0.$$

Exemple 4. Calculer l'intégrale $\oint\limits_L (x + y)\,dy$, où le contour L est le périmètre du rectangle engendré par les droites $x = 0$, $y = 0$, $x = 1$ et $y = 1$ (fig. 183).

Solution. Le sens direct est indiqué sur la figure 183 par des flèches. Partageons le contour d'intégration L en quatre parties :

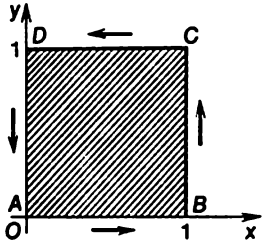


Fig. 183

$$\oint\limits_L = \int\limits_{\overline{AB}} + \int\limits_{\overline{BC}} + \int\limits_{\overline{CD}} + \int\limits_{\overline{DA}}.$$

Il est immédiat que les intégrales le long de AB et CD sont nulles, puisque y est constant sur ces segments, donc $dy = 0$. Reste donc à calculer les intégrales le long de BC et DA. En remplaçant x par y et $y(x)$ par $x(y)$ dans la première formule (14), on obtient

$$\int\limits_{\overline{BC}} (x+y)\,dy = \int\limits_0^1 (1+y)\,dy = \left[y + \frac{y^2}{2}\right]_0^1 = \frac{3}{2},$$

$$\int\limits_{\overline{DA}} (x+y)\,dy = \int\limits_1^0 (0+y)\,dy = \left[\frac{y^2}{2}\right]_1^0 = -\frac{1}{2}.$$

En définitive,

$$\oint\limits_L (x+y)\,dy = \frac{3}{2} - \frac{1}{2} = 1.$$

Exemple 5. Calculer l'intégrale $\int\limits_{\overset{\frown}{AB}} 3x^2y\,dx + (x^3+1)\,dy$, où

a) AB est le segment de droite $y = x$ d'origine (0; 0) et d'extrémité (1; 1).

b) $\overset{\frown}{AB}$ est le segment de parabole $y = x^2$ reliant ces points;
c) AB est une ligne polygonale passant par les points (0; 0), (1; 0), (1; 1) (fig. 184).

Solution. La troisième formule (14) nous donne

$$\text{a)} \int\limits_{\overline{AB}} 3x^2 y\,dx + (x^3+1)\,dy = \int\limits_0^1 (4x^3+1)\,dx = 2;$$

$$\text{b)} \int\limits_{\overset{\frown}{AB}} 3x^2 y\,dx + (x^3+1)\,dy = \int\limits_0^1 (5x^4+2x)\,dx = 2;$$

$$\text{c)} \int\limits_{\overline{AB}} 3x^2 y\,dx + (x^3+1)\,dy = \int\limits_0^1 3x^2 \cdot 0\,dx + \int\limits_0^1 (1+1)\,dy =$$

$$= \int\limits_0^1 2dy = 2.$$

On remarquera que le résultat de l'intégration ne dépend pas du chemin suivi pour aller de A en B. L'explication de ce fait sera donnée au § 7.

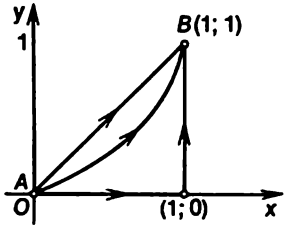


Fig. 184

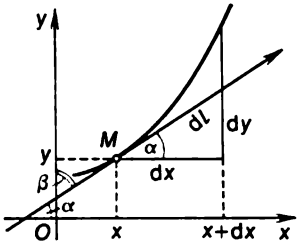


Fig. 185

5. **Relation entre intégrales curvilignes de première et de seconde espèce.** Désignons par α et β les angles formés avec les axes de coordonnées par la tangente orientée *) à la courbe AB en $M(x; y)$ (fig. 185); on obtient alors les relations

$$dx = \cos\alpha\,dl, \quad dy = \cos\beta\,dl \text{ **)}. \tag{15}$$

En portant les expressions (15) de dx et dy dans les intégrales curvilignes de seconde espèce, on obtient les intégrales de première

*) Pour sens direct on prend le sens des t croissants
**) Voir remarque du n° 2.

espèce :

$$\int_{\overset{\frown}{AB}} P(x, y)\,dx = \int_{\overset{\frown}{AB}} P(x, y)\cos\alpha\,dl,$$
$$\int_{\overset{\frown}{AB}} Q(x, y)\,dy = \int_{\overset{\frown}{AB}} Q(x, y)\cos\beta\,dl, \tag{16}$$
$$\int_{\overset{\frown}{AB}} P(x, y)\,dx + Q(x, y)\,dy = \int_{\overset{\frown}{AB}} [P(x, y)\cos\alpha + Q(x, y)\cos\beta]\,dl.$$

Donc les formules (16) expriment les intégrales curvilignes de seconde espèce par l'intermédiaire d'intégrales de première espèce et établissent ainsi un lien entre elles. Lorsque le point courant parcourt la courbe dans le sens opposé, cos α, cos β, dx et dy changent de signe et les formules (16) restent en vigueur.

Signalons en conclusion que nous avons étudié les intégrales curvilignes pour des courbes planes. Il est possible de généraliser sans peine leurs définitions et propriétés aux courbes gauches.

Soit AB une courbe gauche sur laquelle sont définies des fonctions $f(x, y, z)$, $P(x, y, z)$, $Q(x, y, z)$ et $R(x, y, z)$. On peut comme dans le cas d'une courbe plane définir les intégrales curvilignes de première et de seconde espèce suivantes :

$$\int_{\overset{\frown}{AB}} f(x, y, z)\,dl, \int_{\overset{\frown}{AB}} P(x, y, z)\,dx, \int_{\overset{\frown}{AB}} Q(x, y, z)\,dy, \int_{\overset{\frown}{AB}} R(x, y, z)\,dz,$$
$$\int_{\overset{\frown}{AB}} P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz.$$

La technique de calcul de ces intégrales est à peu de choses près la même que pour les intégrales le long d'une courbe plane.

§ 6. Formule de Green

La formule de Green *) établit un lien entre les intégrales curvilignes et les intégrales doubles. Elle est souvent sollicitée en analyse et ses applications.

Prouvons cette formule pour un domaine fermé dont le contour n'est pas coupé en plus de deux points par des droites parallèles aux axes de coordonnées. Pour simplifier, ces domaines seront dits simples. On admet que le contour est différentiable ou différentiable par morceaux.

Théorème 13.6. *Soit G un domaine fermé simple de contour L et soient $P(x, y)$ et $Q(x, y)$ des fonctions continues avec leurs déri-*

*) Green Georges (1793-1841), mathématicien et physicien anglais.

vées partielles $\frac{\partial P}{\partial y}$ *et* $\frac{\partial Q}{\partial x}$ *sur ce domaine. On a alors la formule*

$$\iint_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy = \oint_L P\, dx + Q\, dy, \tag{1}$$

dite formule de Green.

Démonstration. Supposons que le contour L du domaine G est définissable aussi bien par des équations $x = x_1(y)$, $x = x_2(y)$ $(c \leqslant y \leqslant d)$, $x_1(y) \leqslant x_2(y)$, que par des équations $y = y_1(x)$, $y = y_2(x)$ $(a \leqslant x \leqslant b)$, $y_1(x) \leqslant y_2(x)$ (fig. 186). Considérons d'abord un domaine G défini par les doubles inégalités $a \leqslant x \leqslant b$, $y_1(x) \leqslant y \leqslant y_2(x)$ et transformons l'intégrale double

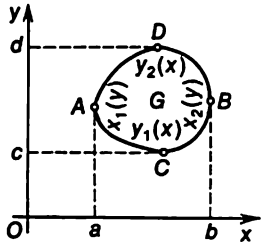


Fig. 186

$$\iint_G \frac{\partial P}{\partial y}\, dx\, dy$$

en intégrale curviligne. A cet effet, réduisons-la à une intégrale itérée et intégrons par rapport à y à l'aide de la formule de Newton-Leibniz. On obtient

$$\iint_G \frac{\partial P}{\partial y}\, dx\, dy = \int_a^b dx \int_{y_1(x)}^{y_2(x)} \frac{\partial P}{\partial y}\, dy =$$

$$= \int_a^b [P(x,\ y_2(x)) - P(x,\ y_1(x))]\, dx =$$

$$= \int_a^b P(x,\ y_2(x))\, dx - \int_a^b P(x,\ y_1(x))\, dx.$$

Chacune de ces deux intégrales définies est égale à l'intégrale curviligne de seconde espèce le long de la courbe correspondante (cf. formule (14), § 5), plus exactement,

$$\int_a^b P(x,\ y_2(x))\, dx = \int_{\overset{\frown}{ADB}} P(x,\ y)\, dx = -\int_{\overset{\frown}{BDA}} P(x,\ y)\, dx,$$

$$\int_a^b P(x,\ y_1(x))\, dx = \int_{\overset{\frown}{ACB}} P(x,\ y)\, dx.$$

Donc

$$\iint_G \frac{\partial P}{\partial y}\,dx\,dy = -\Big[\int_{\overset{\frown}{BDA}} P(x,\ y)\,dx + \int_{\overset{\frown}{ACB}} P(x,\ y)\,dx\Big],$$

c'est-à-dire que

$$\iint_G \frac{\partial P}{\partial y}\,dx\,dy = -\oint_L P(x,\ y)\,dx. \tag{2}$$

On démontre de façon analogue la formule

$$\iint_G \frac{\partial Q}{\partial x}\,dx\,dy = \oint_L Q(x,\ y)\,dy, \tag{3}$$

le domaine G étant défini ici par les doubles inégalités $c \leqslant y \leqslant d$, $x_1(y) \leqslant x \leqslant x_2(y)$.

En retranchant terme à terme l'égalité (3) de (2), on obtient la formule (1) annoncée. ■

Remarque. La formule de Green reste valable pour tout domaine fermé G divisible en un nombre fini de domaines fermés simples par adjonction de lignes auxiliaires. Supposons en effet que le domaine G de contour L est de la forme représentée sur la figure 187. Divisons-le en deux domaines simples: G_1 et G_2, justiciables chacun de la formule (1). Ecrivons séparément la formule de Green pour G_1 et G_2 et additionnons les égalités obtenues. Au premier membre on aura une intégrale double étendue au domaine G tout entier, au second, une intégrale curviligne le long du contour L de G, puisque l'intégrale curviligne le long de la courbe auxiliaire est prise une fois dans un sens et une autre fois dans le sens opposé.

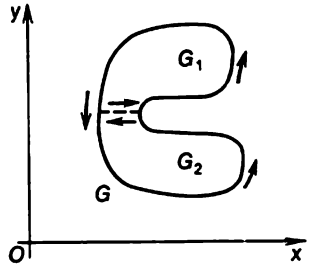


Fig. 187

Exemple. Appliquer la formule de Green au calcul de l'intégrale curviligne $\oint_L (x - y)\,dx + (x + y)\,dy$, où L est le cercle $x^2 + y^2 = R^2$.

Solution. Les fonctions $P(x,\ y) = x - y$, $Q(x,\ y) = x + y$ et $\frac{\partial P}{\partial y} = -1$, $\frac{\partial Q}{\partial x} = 1$ sont continues sur le disque fermé $x^2 + y^2 \leqslant R^2$. Donc cette intégrale est justiciable de la formule de

Green d'après le théorème 13.6. On a

$$\oint_L (x-y)\,dx+(x+y)\,dy=\iint_G [1-(-1)]\,dx\,dy=$$

$$=2\iint_G dx\,dy=2s=2\pi R^2.$$

Remarquons que le résultat obtenu se vérifie par un calcul direct de cette intégrale.

§ 7. Conditions d'indépendance de l'intégrale curviligne par rapport au chemin d'intégration

On a vu en résolvant l'exemple 5 (cf. § 5, n° 4) que dans certains cas la valeur de l'intégrale curviligne $\int_{\widehat{AB}} P\,dx+Q\,dy$ ne dépend pas du chemin d'intégration mais seulement de ses extrémités A et B. Etablissons les conditions de cette indépendance. La formule de Green joue un rôle important dans l'étude de cette question.

Précisons la nature des domaines envisagés plus bas.

Définition. On dit qu'un domaine plan G est *simplement connexe* si pour tout contour fermé L situé dans G, l'intérieur de L est entièrement compris dans G.

De façon plus imagée, un domaine simplement connexe est un domaine ne présentant pas de trous. Comme exemples de domaines simplement connexes citons l'intérieur d'un cercle, d'une ellipse, d'un polygone, etc. Un exemple élémentaire de domaine non simplement connexe est la couronne limitée par les cercles $x^2+y^2=1$ et $x^2+y^2=3$. En effet, le cercle $x^2+y^2=2$ situé dans ce domaine contient des points qui n'appartiennent pas à cette couronne, par exemple l'origine des coordonnées (0 ; 0).

Théorème 13.7. *Soient $P(x, y)$ et $Q(x, y)$ des fonctions définies et continues avec leurs dérivées partielles $\frac{\partial P}{\partial y}$ et $\frac{\partial Q}{\partial x}$ sur un domaine fermé simplement connexe G. Les quatre conditions suivantes sont alors équivalentes, c'est-à-dire que l'une d'elles entraîne les trois autres :*

1) *pour toute courbe fermée L différentiable par morceaux située dans G,*

$$\oint_L P\,dx+Q\,dy=0 ;$$

2) *pour tout couple de points A et B du domaine G la valeur de l'intégrale*

$$\int_{\overset{\frown}{AB}} P\,dx+Q\,dy$$

ne dépend pas du chemin suivi à l'intérieur de G pour aller de A en B;

3) *l'expression* $P\,dx + Q\,dy$ *est la différentielle totale* *) *d'une fonction définie sur G. En d'autres termes, il existe une fonction* $F(x, y)$ *définie sur G telle que*

$$dF = P\,dx+Q\,dy;$$

4) *sur le domaine G*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}. \qquad (1)$$

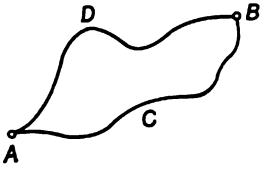


Fig. 188

Démonstration. Prouvons ce théorème dans l'ordre suivant:

$$1 \to 2 \to 3 \to 4 \to 1,$$

c'est-à-dire montrons que la première condition entraîne la seconde, la seconde entraîne la troisième, la troisième entraîne la quatrième et enfin la quatrième entraîne la première. Ceci prouvera l'équivalence de ces conditions.

Première étape. $1 \to 2$. Relions deux points A et B du domaine G par deux courbes ACB et ADB différentiables par morceaux (fig. 188). La réunion de ces courbes est le contour $L = ACB + BDA$ situé dans G. D'après la condition 1),

$$\oint_L P\,dx+Q\,dy=0,$$

mais

$$\oint_L P\,dx+Q\,dy=\int_{\overset{\frown}{ABC}} P\,dx+Q\,dy+\int_{\overset{\frown}{BDA}} P\,dx+Q\,dy=$$

$$=\int_{\overset{\frown}{ACB}} P\,dx+Q\,dy-\int_{\overset{\frown}{ADB}} P\,dx+Q\,dy,$$

donc

$$\int_{\overset{\frown}{ACB}} P\,dx+Q\,dy=\int_{\overset{\frown}{ADB}} P\,dx+Q\,dy,$$

ce qu'on voulait.

*) En théorie des intégrales curvilignes de seconde espèce, la différentielle de la fonction $F(x, y)$ est généralement appelée *différentielle totale*.

Deuxième étape. 2→3. Supposons que l'intégrale $\int\limits_{\overset{\frown}{AB}} P\,dx+Q\,dy$ ne dépend pas du chemin suivi pour aller de A en B. Si l'on fixe le point A, c'est-à-dire si $A = A(x_0; y_0)$ cette intégrale sera une fonction des coordonnées x et y du point $B = B(x; y)$:

$$\int\limits_{\overset{\frown}{AB}} P\,dx+Q\,dy = F(x,\ y).$$

Montrons que la fonction $F(x, y)$ est différentiable et que

$$dF = P\,dx + Q\,dy. \tag{2}$$

Il suffit de prouver pour cela qu'en chaque point B de G existent les dérivées partielles $\frac{\partial F}{\partial x}$ et $\frac{\partial F}{\partial y}$ et de plus que

$$\frac{\partial F}{\partial x} = P(x,\ y), \quad \frac{\partial F}{\partial y} = Q(x,\ y). \tag{3}$$

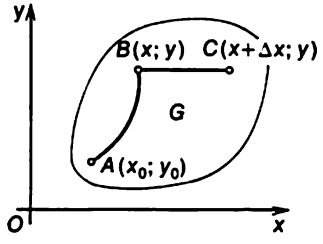


Fig. 189

Les fonctions $P(x, y)$ et $Q(x, y)$ étant continues dans G, les relations (3) entraînent la différentiabilité de $F(x, y)$ et l'égalité (2).

Pour prouver l'existence de $\frac{\partial F}{\partial x}$ et la première égalité (3), formons l'accroissement partiel par rapport à x de $F(x, y)$ au point $B(x; y)$:

$$\Delta_x F = F(x+\Delta x,\ y) - F(x,\ y) =$$

$$= \int\limits_{\overset{\frown}{AC}} P\,dx+Q\,dy - \int\limits_{\overset{\frown}{AB}} P\,dx+Q\,dy = \int\limits_{\overline{BC}} P\,dx+Q\,dy,$$

où $C = C(x + \Delta x; y)$ (fig. 189). L'intégrale étant par hypothèse indépendante de la forme du chemin d'intégration, prenons pour chemin le segment de droite BC. Alors

$$\Delta_x F = \int\limits_{\overline{BC}} P\,dx+Q\,dy = \int\limits_{\overline{BC}} P\,dx = \int\limits_{x}^{x+\Delta x} P(x,\ y)\,dx.$$

En appliquant le théorème de la moyenne à la dernière intégrale, on trouve

$$\Delta_x F = P(x + \theta\,\Delta x,\ y)\,\Delta x, \quad 0 \leqslant \theta < 1,$$

d'où

$$\frac{\Delta_x F}{\Delta x} = P(x+\theta\Delta x,\ y) \quad 0 < \theta < 1.$$

Donc

$$\frac{\partial F}{\partial x} = \lim_{\Delta x \to 0} P(x + \theta \Delta x,\ y) = P(x,\ y),$$

puisque $P(x, y)$ est continue par hypothèse. On démontre par analogie que $\frac{\partial F}{\partial y} = Q(x, y)$. On a ainsi établi la condition 3).

Troisième étape. $3 \to 4$. Supposons que sur le domaine G est définie une fonction $F(x, y)$ telle que $dF = P\,dx + Q\,dy$. Alors

$$\frac{\partial F}{\partial x} = P, \quad \frac{\partial F}{\partial y} = Q,$$

et d'après le théorème relatif à l'égalité des dérivées mixtes,

$$\frac{\partial Q}{\partial x} = \frac{\partial^2 F}{\partial x\,\partial y} = \frac{\partial^2 F}{\partial y\,\partial x} = \frac{\partial P}{\partial y},$$

c'est-à-dire qu'on a obtenu l'égalité annoncée (1).

Quatrième étape. $4 \to 1$. Supposons que la condition 4) est remplie et soit L une courbe différentiable par morceaux située dans G et limitant un domaine G^*. Alors en appliquant la formule de Green au domaine G^* (on se sert ici de la connexité simple de G), on obtient

$$\oint_L P\,dx + Q\,dy = \iint_{G^*} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy.$$

L'intégrale du second membre est nulle d'après la condition 4). Donc

$$\oint_L P\,dx + Q\,dy = 0$$

pour tout contour L situé dans G. ■

Remarque. De l'équivalence des conditions 1) à 4) du théorème 13.7 il s'ensuit, en particulier, que la condition 3) est une condition nécessaire et suffisante pour que l'intégrale curviligne ne dépende pas du chemin d'intégration. Mais la condition 4) est une condition nécessaire et suffisante qui est plus commode dans les applications.

Le théorème 13.7 permet de dire facilement si l'intégrale curviligne dépend ou non du chemin d'intégration. Ainsi, par exemple, $\int_{\widehat{AB}} e^y\,dx - y\,dy$ dépend du chemin d'intégration dans tout domaine, puisque $\frac{\partial P}{\partial y} = e^y \neq 0 = \frac{\partial Q}{\partial x}$. Mettons l'accent sur le fait que toutes les conditions sont essentielles. Considé-

rons par exemple l'intégrale

$$I = \oint_L \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy,$$

où L est un cercle de rayon R centré en (0 ; 0). On a

$$P = -\frac{y}{x^2+y^2}, \quad \frac{\partial P}{\partial y} = \frac{-x^2-y^2+2y^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2},$$

$$Q = \frac{x}{x^2+y^2}, \quad \frac{\partial Q}{\partial x} = \frac{x^2+y^2-2x^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}.$$

On voit que la condition d'indépendance par rapport au chemin d'intégration est formellement remplie, et pourtant l'intégrale le long de L n'est pas nulle. En effet, si l'on définit le cercle par les équations $x = R\cos t$, $y = R\sin t$, on obtient

$$I = \int_0^{2\pi} \frac{-R\sin t\,(-R\sin t) + R\cos t\,R\cos t}{R^2\cos^2 t + R^2\sin^2 t}\,dt = \int_0^{2\pi} dt = 2\pi.$$

En fait, il n'y a aucune contradiction avec le théorème. Tout simplement l'une des conditions du théorème n'est pas remplie : les fonctions P et Q et leurs dérivées partielles $\frac{\partial P}{\partial y}$ et $\frac{\partial Q}{\partial x}$ ne sont pas définies en (0 ; 0) ; quant au disque limité par L et privé du point (0 ; 0), il n'est plus un domaine simplement connexe (l'origine des coordonnées joue le rôle de « trou »).

§ 8. Intégration des différentielles totales

L'étude des conditions d'indépendance de l'intégrale curviligne $\int_{\widehat{AB}} P\,dx + Q\,dy$ par rapport au choix du chemin d'intégration nous permet de résoudre directement le problème de l'intégration des différentielles totales et de déterminer une fonction d'après sa différentielle totale.

On a montré que si des fonctions $P(x, y)$ et $Q(x, y)$ et leurs dérivées partielles $\frac{\partial P}{\partial y}$ et $\frac{\partial Q}{\partial x}$ étaient continues dans un domaine fermé G, alors l'expression

$$P\,dx + Q\,dy \tag{1}$$

est la différentielle totale d'une fonction sur ce domaine si et seulement si

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

Nous avons prouvé par ailleurs que si cette égalité avait lieu, la condition

$$dF = P\,dx + Q\,dy$$

serait vérifiée par la fonction

$$F(x, y) = \int\limits_{\overset{\frown}{AB}} P\,dx + Q\,dy = \int\limits_{(x_0;\ y_0)}^{(x;\ y)} P\,dx + Q\,dy. \qquad (2)$$

Supposons maintenant que l'expression (1) est la différentielle totale d'une fonction $\Phi(x, y)$. Alors $\frac{\partial \Phi}{\partial x} = P$, $\frac{\partial \Phi}{\partial y} = Q$ et la différence $\Phi(x, y) - F(x, y)$ (cf. remarque suivant le théorème 12.6) est une quantité constante. Donc

$$\Phi(x, y) = F(x, y) + C, \qquad (3)$$

où C est une constante. En posant $x = x_0$, $y = y_0$, de (2) on tire $F(x_0, y_0) = 0$ et de (3), la valeur de C: $C = \Phi(x_0, y_0)$. Maintenant on peut écrire (3) sous la forme

$$F(x, y) = \Phi(x, y) - \Phi(x_0, y_0),$$

et l'égalité (2), sous la forme

$$\int\limits_{(x_0;\ y_0)}^{(x;\ y)} P\,dx + Q\,dy = \Phi(x, y) - \Phi(x_0, y_0).$$

Si enfin on pose $x = x_1$, $y = y_1$, on obtient la formule

$$\int\limits_{(x_0;\ y_0)}^{(x_1;\ y_1)} P\,dx + Q\,dy = \Phi(x_1, y_1) - \Phi(x_0, y_0) = \Phi(x, y)\Big|_{(x_0;\ y_0)}^{(x_1;\ y_1)}. \qquad (4)$$

La formule (4) rappelle la formule de Newton-Leibniz, mais elle n'est valable que si l'intégrale ne dépend pas du choix du chemin d'intégration.

En nous servant des résultats établis, nous sommes en mesure d'indiquer une méthode de remontée à une fonction $F(x, y)$ dont la différentielle totale est donnée par l'expression (1).

La formule

$$F(x, y) = \int\limits_{(x_0;\ y_0)}^{(x;\ y)} P\,dx + Q\,dy + C, \qquad (5)$$

où (x_0, y_0) est un point fixe et C une constante arbitraire, permet de définir toutes les fonctions dont la différentielle totale est l'intégrant.

Pour déterminer une fonction $F(x, y)$ à l'aide de la formule (5), il suffit, après avoir choisi un point arbitraire du domaine G,

de calculer une intégrale curviligne le long de toute courbe reliant les points $(x_0; y_0)$ et $(x; y)$. Puisque dans la formule (5) l'intégrale ne dépend pas du choix du chemin d'intégration, il est commode de prendre pour chemin par exemple une ligne polygonale dont les segments soient parallèles aux axes de coordonnées (fig. 190). Alors

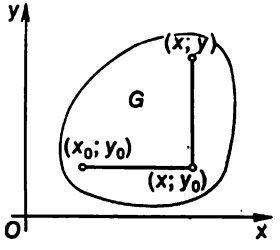


Fig. 190

$$\int\limits_{(x_0;\ y_0)}^{(x;\ y)} P\,dx + Q\,dy = \int\limits_{(x_0;\ y_0)}^{(x;\ y_0)} P\,dx + Q\,dy + \int\limits_{(x;\ y_0)}^{(x;\ y)} P\,dx + Q\,dy.$$

Puisque, d'une part, $y = y_0$ et $dy = 0$ sur le segment $[(x_0; y_0); (x; y_0)]$ et, d'autre part, $dx = 0$ sur le segment $[(x; y_0); (x; y)]$, l'égalité (5) devient

$$F(x, y) = \int\limits_{x_0}^{x} P(x, y_0)\,dx + \int\limits_{y_0}^{y} Q(x, y)\,dy + C,$$

où la première intégrale définie est calculée pour y constant égal à y_0 et la deuxième, pour x constant.

Exemple 1. Vérifier si l'expression $(x^2 + 2xy - y^2)\,dx + (x^2 - 2xy - y^2)\,dy$ est la différentielle totale d'une fonction $F(x, y)$ et, le cas échéant, trouver $F(x, y)$.

Solution. Dans cette expression les fonctions

$$P(x, y) = x^2 + 2xy - y^2,\ Q(x, y) = x^2 - 2xy - y^2 \qquad (6)$$

sont continues avec leurs dérivées partielles $\frac{\partial P}{\partial y} = 2x - 2y$ et $\frac{\partial Q}{\partial x} = 2x - 2y$ qui sont égales. Donc cette expression est la différentielle totale d'une fonction $F(x, y)$. Pour trouver cette fonction, utilisons la formule (2), où $A(x_0; y_0)$ est un point fixe et $B(x; y)$, un point variable.

Pour $A(x_0; y_0)$ il est commode de prendre le point $(0; 0)$. Puisque l'intégrale curviligne $\int\limits_{(0;\ 0)}^{(x;\ y)} (x^2 + 2xy - y^2)\,dx + (x^2 - 2xy - y^2)\,dy$ ne dépend pas du choix du chemin d'intégration, relions les points $(0; 0)$ et $(x; y)$ par une ligne polygonale de segments parallèles aux axes de coordonnées. Il suffit pour cela de

prendre le point $(x; 0)$ (ou le point $(0; y)$) (fig. 191). L'un des segments de la ligne polygonale sera porté par un axe de coordonnées. On a

$$F(x, y) = \int_{(0; 0)}^{(x; y)} (x^2 + 2xy - y^2)\,dx + (x^2 - 2xy - y^2)\,dy + C =$$

$$= \int_0^x x^2\,dx + \int_0^y (x^2 - 2xy - y^2)\,dy + C = \frac{x^3}{3} + x^2y - xy^2 - \frac{y^3}{3} + C,$$

où C est une constante arbitraire.

Quand on cherche à déterminer une fonction par sa différentielle totale, il est commode de procéder comme suit. Si

$$\frac{\partial F}{\partial x} = P, \quad \frac{\partial F}{\partial y} = Q,$$

en intégrant la première égalité par rapport à x, on obtient

$$F(x, y) = \int P\,dx + f_1(y), \qquad (7)$$

et en intégrant la deuxième par rapport à y,

$$F(x, y) = \int Q\,dy + f_2(x), \qquad (8)$$

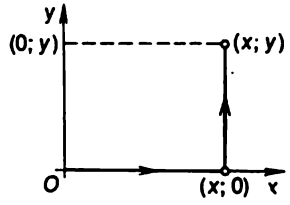


Fig. 191

où $f_1(y)$ et $f_2(x)$ sont des fonctions arbitraires. Si l'on choisit les fonctions $f_1(y)$ et $f_2(x)$ de telle sorte que les seconds membres de (7) et (8) soient confondus, la fonction $F(x, y)$ sera une fonction dont la différentielle totale est $P\,dx + Q\,dy$.

Ainsi, supposons par exemple que $dF = (2xy + 1)\,dx + (x^2 + 3y^2)\,dy$. Intégrons le coefficient de dx par rapport à x:

$$\int (2xy + 1)\,dx = x^2y + x + f_1(y); \qquad (9)$$

et le coefficient de dy par rapport à y:

$$\int (x^2 + 3y^2)\,dy = yx^2 + y^3 + f_2(x). \qquad (10)$$

Les seconds membres de (9) et (10) sont confondus si l'on pose $f_1(y) = y^3 + C$ et $f_2(x) = x + C$.

Donc

$$F(x, y) = yx^2 + y^3 + x + C.$$

Exemple 2. Calculer l'intégrale curviligne $\int_{(-1; 2)}^{(2; 3)} y\,dx + x\,dy$.

Solution. Les fonctions

$$P = y, \quad Q = x, \quad \frac{\partial P}{\partial y} = 1, \quad \frac{\partial Q}{\partial x} = 1$$

sont continues et leurs dérivées partielles sont égales. Donc l'expression $y\,dx + x\,dy$ est la différentielle totale $dF(x, y)$ et cette intégrale ne dépend pas du choix du chemin d'intégration. Les formules (7) et (8) nous donnent $F(x, y) = xy$ et la formule (4),

$$\int_{(-1;\,2)}^{(2;\,3)} y\,dx + x\,dy = xy\Big|_{(-1;\,2)}^{(2;3)} = 2\cdot 3 - (-1)\cdot 2 = 6 + 2 = 8.$$

Remarquons que cette intégrale se calcule directement si par exemple on prend pour chemin d'intégration la ligne polygonale de segments parallèles aux axes de coordonnées passant par les points (—1 ; 2), (2 ; 2) et (2 ; 3) (on laisse ceci au soin du lecteur).

§ 9. Quelques applications des intégrales curvilignes de seconde espèce

Les intégrales curvilignes de seconde espèce possèdent un vaste spectre d'application en géométrie, en physique et en technique. On se bornera à l'étude de deux problèmes : le calcul des aires planes et le calcul du travail d'une force.

1. Calcul d'une aire à l'aide de la formule de Green. Soit G un domaine de contour L et d'aire s. On sait que l'intégrale double $\iint_G f(x, y)\,dx\,dy$ pour $f(x, y) \equiv 1$ exprime l'aire s du domaine G. Si donc dans la formule de Green on choisit les fonctions $P(x, y)$ et $Q(x, y)$ telles que $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \equiv 1$, l'aire s de G sera définie par la formule

$$s = \iint_G dx\,dy = \oint_L P\,dx + Q\,dy.$$

Posons $Q(x, y) = x$ et $P(x, y) = 0$; alors $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$ et

$$s = \oint_L x\,dy.$$

En posant $P(x, y) = -y$ et $Q(x, y) = 0$, on trouve par analogie

$$s = -\oint_L y\,dx,$$

et pour $P(x, y) = -y/2$, et $Q(x, y) = x/2$

$$s = \frac{1}{2} \oint_L x\,dy - y\,dx. \tag{1}$$

Nous avons ainsi obtenu trois formules pour le calcul d'aires planes de contour L.

Exemple 1. Trouver l'aire de l'ellipse.

Solution. Calculons cette aire par exemple à l'aide de la formule (1). L'équation paramétrique de l'ellipse étant $x = a\cos t$, $y = b\sin t$, $0 \leqslant t \leqslant 2\pi$, on a $dx = -a\sin t\,dt$; $dy = b\cos t\,dt$ et d'après la formule (1)

$$s = \frac{1}{2} \oint_l x\,dy - y\,dx =$$

$$= \frac{1}{2} \int_0^{2\pi} (a\cos t\, b\cos t + b\sin t\, a\sin t)\,dt = \frac{ab}{2} \int_0^{2\pi} dt = \pi ab.$$

2. Travail d'une force. On sait que le travail fourni par une force variable $F(x)$ portée par l'axe Ox pour déplacer un point matériel le long de cet axe d'un point $x = a$ à un point $x = b$ $(a < b)$ est donné par l'intégrale définie $A = \int_a^b F(x)\,dx$ (chap. VI, § 8, n° 6). Posons le problème plus général.

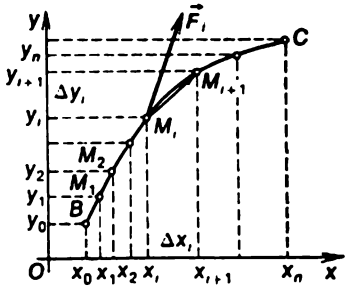


Fig. 192

Supposons qu'un point matériel se déplace le long d'une courbe plane continue BC de B vers C sous l'action d'une force $\vec{F}$. On suppose que la force $\vec{F}$ est une variable dépendant de la position du point matériel sur la courbe BC. Calculons le travail fourni par la force $\vec{F}$ pour déplacer le point matériel de B en C. A cet effet, partageons la courbe BC en n parties par les points $B = M_0$, M_1, M_2, ..., M_i, M_{i+1}, ..., $M_n = C$ (fig. 192). Remplaçons la force $\vec{F}$ sur $\overset{\frown}{M_iM_{i+1}}$ par une valeur constante égale à sa valeur en M_i, c'est-à-dire que $\vec{F}_i \approx \vec{F}(M_i)$, et le mouvement du point matériel le long de l'arc $\overset{\frown}{M_iM_{i+1}}$, par le mouvement sur le segment M_iM_{i+1}. Le travail fourni par la force constante $\vec{F}(M_i)$ le long du segment M_iM_{i+1} peut être pris pour approxi-

mation du travail A_i de la force variable $\vec{F}$ le long de l'arc $\overset{\frown}{M_iM_{i+1}}$, c'est-à-dire que

$$A_i \approx \vec{F}(M_i) \cdot \overrightarrow{M_iM_{i+1}}.$$

Le second membre de cette relation approchée est le produit scalaire des vecteurs $\vec{F}(M_i)$ et $\overrightarrow{M_iM_{i+1}}$. Il est égal à la somme des produits deux à deux des composantes correspondantes de ces vecteurs, c'est-à-dire que si $\vec{F}(M_i) = \{P(M_i);\ Q(M_i)\}$, $\overrightarrow{M_iM_{i+1}} = \{\Delta x_i;\ \Delta y_i\}$, alors $A_i \approx P(M_i)\,\Delta x_i + Q(M_i)\,\Delta y_i$. En sommant sur i entre 1 et n, on obtient la valeur approchée du travail A le long de la courbe BC :

$$A \approx \sum_{i=1}^{n} (P(M_i)\,\Delta x_i + Q(M_i)\,\Delta y_i) =$$

$$= \sum_{i=1}^{n} P(M_i)\,\Delta x_i + \sum_{i=1}^{n} Q(M_i)\,\Delta y_i. \qquad (2)$$

Pour valeur exacte du travail A on prend la limite vers laquelle tend cette valeur approchée lorsque le plus grand des arcs $\overset{\frown}{M_iM_{i+1}}$ tend vers 0. Mais, d'autre part, l'expression (2) est une somme de deux sommes intégrales pour les fonctions $P(x, y)$ et $Q(x, y)$ définies sur la courbe BC. La limite de cette somme est par définition une intégrale curviligne de seconde espèce. Donc le travail fourni par la force est défini par la formule

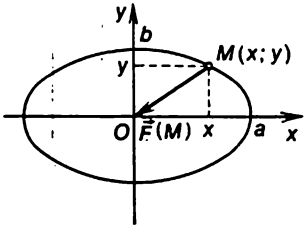


Fig. 193

$$A = \int_{\overset{\frown}{BC}} P\,dx + Q\,dy, \qquad (3)$$

où P et Q sont les composantes de la force $\vec{F}$.

Si l'on pose ce problème dans l'espace, sa résolution se ramène au calcul d'une intégrale curviligne de seconde espèce le long d'une courbe gauche à l'aide de la formule

$$A = \int_{\overset{\frown}{BC}} P\,dx + Q\,dy + R\,dz.$$

Exemple 2. Calculer le travail fourni par une force $\vec{F}(x, y)$ pour déplacer un point matériel le long d'une ellipse de centre (0 ; 0) dans le sens direct sachant qu'en chaque point $(x; y)$ de l'ellipse cette force est dirigée vers (0 ; 0) et sa mesure est égale à la distance de $(x; y)$ au centre (0 ; 0) (fig. 193).

Solution. Par hypothèse, $|\vec{F}(x, y)| = \sqrt{x^2 + y^2}$; les composantes de la force $\vec{F}$ sont $P = -x$ et $Q = -y$ (le signe « — » provient de ce que la force est dirigée vers le centre (0; 0)). La formule (3) nous donne $A = -\oint\limits_L x\,dx + y\,dy$, où L est l'ellipse $x = a\cos t$, $y = b\sin t$, $0 \leqslant t \leqslant 2\pi$. Donc

$$A = \frac{a^2 - b^2}{2} \int\limits_0^{2\pi} \sin 2t\,dt = \frac{a^2 - b^2}{4}(-\cos 2t)\Big|_0^{2\pi} = 0.$$

Remarquons que la nullité de l'intégrale est due au fait que l'intégrant est la différentielle totale d'une fonction (on demande au lecteur de trouver cette fonction).

§ 10. Intégrales triples

Au début de ce chapitre nous avons introduit la notion d'intégrale double d'une fonction de deux variables. Définissons maintenant la notion d'intégrale triple d'une fonction de trois variables. Les intégrales triples sont largement utilisées dans la résolution de nombreux problèmes de physique et de géométrie.

1. Définition de l'intégrale triple. L'intégrale triple pour une fonction de trois variables se définit comme l'intégrale double pour une fonction de deux variables.

Soit donnée une fonction bornée $f(M) = f(x, y, z)$ sur un domaine fermé borné V de l'espace. Partageons le domaine V en parties sans points intérieurs communs, de volumes $\Delta v_1, \Delta v_2, \ldots, \Delta v_n$. Dans chaque partie prenons un point arbitraire $M_i(\xi_i; \eta_i; \zeta_i)$ et formons la somme

$$\sum_{i=1}^{n} f(\xi_i; \eta_i; \zeta_i)\,\Delta v_i, \tag{1}$$

qui s'appelle *somme intégrale* pour la fonction $f(x, y, z)$ sur le domaine V. Désignons par λ le plus grand diamètre des domaines partiels.

Définition. La limite I, si elle existe, de la somme intégrale (1) lorsque $\lambda \to 0$ s'appelle *intégrale triple* de la fonction $f(x, y, z)$ étendue au domaine V et se représente par l'une des notations suivantes:

$$I = \int\int\int_V f(x, y, z)\,dv = \int\int\int_V f(x, y, z)\,dx\,dy\,dz.$$

Alors la fonction $f(x, y, z)$ est dite intégrable sur le domaine V; V est le domaine d'intégration; x, y et z, les variables d'intégration; dv (ou $dx\,dy\,dz$), l'élément de volume.

Puisque les résultats acquis pour les intégrales doubles et leurs démonstrations se généralisent aux intégrales triples on se contentera simplement de les formuler en les assortissant de brefs éclaircissements.

Les intégrales triples sont une généralisation immédiate des égrales doubles au cas d'un espace à trois dimensions. Elles sont justiciables de conditions nécessaires et suffisantes d'existence et de propriétés identiques à celles des intégrales doubles. Si l'on pose $f(x, y, z) \equiv 1$ sur V, la définition de l'intégrale triple nous donne une formule de calcul du volume de V:

$$v = \iiint_V dv = \iiint_V dx\,dy\,dz.$$

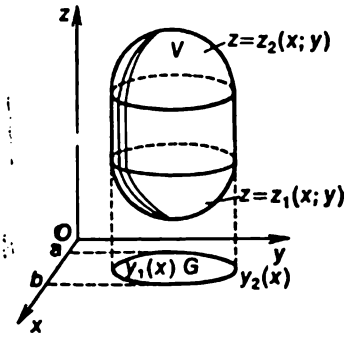


Fig. 194

2. Calcul des intégrales triples. De même que pour les intégrales doubles, le calcul des intégrales triples se ramène au calcul d'intégrales de moindre dimension.

Soit un domaine V limité par les surfaces $z = z_1(x, y)$ et $z = z_2(x, y)$ et latéralement par une surface cylindrique, et soit G le projeté du domaine V sur le plan Oxy (fig. 194). Supposons que les fonctions $z_1(x, y)$ et $z_2(x, y)$ sont définies et continues sur G. Supposons enfin que chaque droite parallèle à Oz coupe le contour de V en deux points au plus. Pour toute fonction $f(x, y, z)$ continue sur V, on a alors la formule

$$\iiint_V f(x, y, z)\,dx\,dy\,dz = \iint_G dx\,dy \int_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\,dz,$$

qui permet de ramener le calcul de l'intégrale triple au calcul successif d'une intégrale définie intérieure par rapport à z (pour x et y constants) et d'une intégrable double extérieure étendue au domaine G.

L'expression

$$I(x, y) = \int_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\,dz$$

est une fonction de deux variables. Si les conditions du théorème 13.4 sont remplies pour cette fonction et pour le domaine

d'intégration G, en passant de l'intégrale double $\iint\limits_G I(x, y)\,dx\,dy$ à l'intégrale itérée, on obtient la formule

$$\iiint\limits_V f(x, y, z)\,dx\,dy\,dz = \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} dy \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\,dz, \qquad (2)$$

qui ramène le calcul de l'intégrale triple au calcul successif de trois intégrales définies. L'ordre d'intégration importe peu, c'est-à-dire que les variables x, y et z peuvent changer de rôle dans la formule (2).

Si, en particulier, V est un parallélépipède de faces $x = a$, $x = b$ $(a < b)$, $y = c$, $y = d$ $(c < d)$, $z = k$, $z = l$ $(k < l)$, la formule (2) devient

$$\iiint\limits_V f(x, y, z)\,dx\,dy\,dz =$$

$$= \int\limits_a^b dx \int\limits_c^d dy \int\limits_k^l f(x, y, z)\,dz. \qquad (3)$$

Dans ce cas l'intégration peut être effectuée dans n'importe quel ordre.

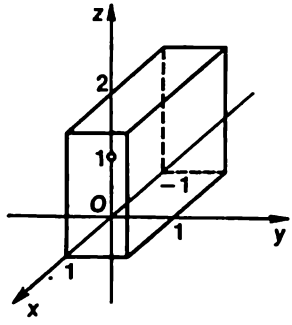


Fig. 195

Exemple 1. Calculer l'intégrale $\iiint\limits_V (x+y-z)\,dx\,dy\,dz$, où V est le parallélépipède limité par les plans $x = -1$, $x = 1$, $y = 0$, $y = 1$, $z = 0$, $z = 2$ (fig. 195).

Solution. La formule (3) nous donne

$$\iiint\limits_V (x+y-z)\,dx\,dy\,dz = \int\limits_{-1}^{1} dx \int\limits_0^1 dy \int\limits_0^2 (x+y-z)\,dz =$$

$$= \int\limits_{-1}^{1} dx \int\limits_0^1 \left[xz+yz-\frac{z^2}{2}\right]_0^2 dy = \int\limits_{-1}^{1} dx \int\limits_0^1 (2x+2y-2)\,dy =$$

$$= \int\limits_{-1}^{1} [2xy+y^2-2y]_0^1\,dx = \int\limits_{-1}^{1} (2x-1)\,dx = [x^2-x]_{-1}^{1} = -2.$$

Exemple 2. Calculer l'intégrale $\iiint\limits_V (x+y+z)\,dx\,dy\,dz$, où V est la pyramide limitée par le plan $x + y + z = 1$ et les plans de coordonnées $x = 0$, $y = 0$ et $z = 0$ (fig. 196).

S o l u t i o n. Le projeté de V sur le plan Oxy est le triangle G limité par les droites $x = 0$, $y = 0$, $y = 1 - x$. La formule (2) nous donne

$$\iiint_V (x+y+z)\,dx\,dy\,dz = \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} (x+y+z)\,dz =$$

$$= \int_0^1 dx \int_0^{1-x} \left[xz+yz+\frac{z^2}{2}\right]_0^{1-x-y} dy =$$

$$= \frac{1}{2}\int_0^1 \left[y - yx^2 - xy^2 - \frac{y^3}{3}\right]_0^{1-x} dx = \frac{1}{6}\int_0^1 (2-3x+x^3)\,dx =$$

$$= \frac{1}{6}\left[2x - \frac{3x^2}{2} + \frac{x^4}{4}\right]_0^1 = \frac{1}{6}\cdot\frac{3}{4} = \frac{1}{8}.$$

3. Changement de variables dans l'intégrale triple. De même que pour l'intégrale double, il existe des formules qui permettent dans l'intégrale triple de passer des coordonnées rectangulaires à d'autres coordonnées et notamment aux coordonnées cylindriques et sphériques qui sont les plus usitées.

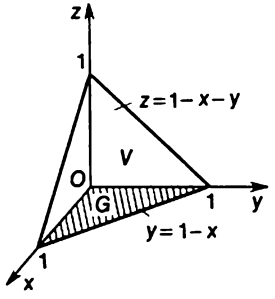


Fig. 196

Le changement de variables dans une intégrale triple s'effectue de la manière suivante.

Si un domaine fermé borné V de l'espace (x, y, z) est appliqué bijectivement sur un domaine V^* de l'espace (u, v, w) par des fonctions continûment dérivables $x = x(u, v, w)$, $y = y(u, v, w)$, $z = z(u, v, w)$ et si le jacobien J

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix} \neq 0$$

sur le domaine V^*, on a la formule

$$\iiint_V f(x, y, z)\,dx\,dy\,dz =$$

$$= \iiint_{V^*} f[x(u, v, w), y(u, v, w), z(u, v, w)]\,|J|\,du\,dv\,dw.$$

En particulier, lorsqu'on passe des coordonnées rectangulaires x, y, z aux coordonnées cylindriques ρ, φ et z (fig. 197) à l'aide des formules

$$x=\rho\cos\varphi,\quad y=\rho\sin\varphi,\quad z=z$$
$$(0\leqslant\rho<+\infty,\quad 0\leqslant\varphi\leqslant 2\pi,\quad -\infty<z<+\infty),$$

le jacobien $J=\rho$ et par suite

$$\iiint\limits_V f(x, y, z)\,dx\,dy\,dz=\iiint\limits_{V^*} f(\rho\cos\varphi,\ \rho\sin\varphi,\ z)\,\rho\,d\rho\,d\varphi\,dz. \quad (4)$$

Ces coordonnées sont dites *cylindriques*, car la surface de coordonnées ρ = const est un cylindre dont les génératrices sont parallèles à l'axe Oz.

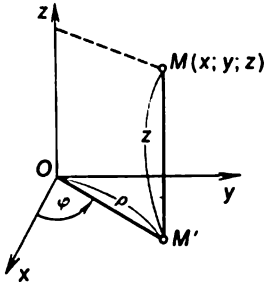


Fig. 197

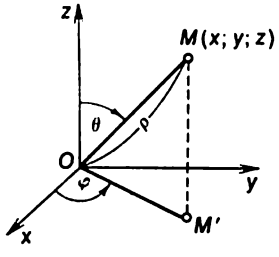


Fig. 198

Le passage des coordonnées rectangulaires x, y et z aux coordonnées sphériques ρ, φ et θ (fig. 198) est réalisé par les formules

$$x=\rho\sin\theta\cos\varphi,\ y=\rho\sin\theta\sin\varphi,\ z=\rho\cos\theta$$
$$(0\leqslant\rho<+\infty,\ 0\leqslant\varphi\leqslant 2\pi,\ 0\leqslant\theta\leqslant\pi);$$

puisque le jacobien $J=\rho^2\sin\theta$, on a

$$\iiint\limits_V f(x, y, z)\,dx\,dy\,dz=$$

$$=\iiint\limits_{V^*} f[\rho\sin\theta\cos\varphi,\ \rho\sin\theta\sin\varphi,\ \rho\cos\theta]\,\rho^2\sin\theta\,d\rho\,d\varphi\,d\theta. \quad (5)$$

Ces coordonnées s'appellent *sphériques*, car la surface de coordonnées ρ = const est une sphère. Les coordonnées sphériques sont parfois appelées *coordonnées polaires dans l'espace.*

Lorsqu'on calcule une intégrale triple en passant aux coordonnées cylindriques ou sphériques, on ne représente pas en général le domaine V^*, on indique les limites d'intégration en s'inspi-

rant directement de la forme du domaine V et en utilisant la signification géométrique des nouvelles coordonnées.

Exemple 3. Calculer l'intégrale $\iiint_V (x^2+y^2)\,dx\,dy\,dz$ en passant aux coordonnées cylindriques $x = \rho\cos\varphi$, $y = \rho\sin\varphi$, $z = z$, où V est le domaine limité par les surfaces $z = x^2 + y^2$ et $z = 1$ (fig. 199).

Solution. Puisque le projeté du domaine V sur le plan Oxy est le disque $x^2 + y^2 \leqslant 1$, la coordonnée φ varie entre 0 et

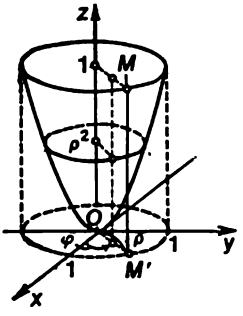


Fig. 199

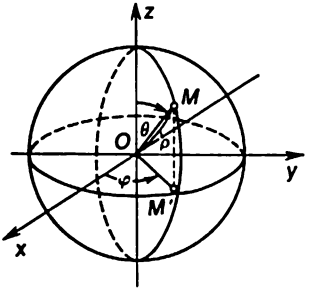


Fig. 200

2π et la coordonnée ρ entre 0 et 1. A toute valeur constante ρ ($0 \leqslant \rho \leqslant 1$) correspond le cylindre $x^2 + y^2 = \rho^2$. En considérant l'intersection de ce cylindre et du domaine V, on trouve que la coordonnée z varie entre les valeurs des points situés sur le paraboloïde $z = x^2 + y^2$ et les valeurs des points situés sur le plan $z = 1$, c'est-à-dire entre $z = \rho^2$ et $z = 1$. En appliquant la formule (4), on obtient

$$\iiint_V (x^2+y^2)\,dx\,dy\,dz = \int_0^{2\pi} d\varphi \int_0^1 d\rho \int_\rho^1 \rho^2\cdot\rho\,dz = \int_0^{2\pi} d\varphi \int_0^1 [\rho^3 z]_\rho^1\,d\rho =$$

$$= \int_0^{2\pi} \left[\frac{\rho^4}{4} - \frac{\rho^6}{6}\right]_0^1 d\varphi = \frac{1}{12}\int_0^{2\pi} d\varphi = \frac{1}{12}\varphi\Big|_0^{2\pi} = \frac{\pi}{6}.$$

Il est difficile de prodiguer des recommandations générales quant au choix du système de coordonnées. Ce choix dépend du domaine d'intégration et de la forme de l'intégrant. Signalons toutefois qu'il est plus commode de se servir de la formule (5) lorsque $f(x, y, z)$ est de la forme $f(x^2 + y^2 + z^2)$ ou que le domaine V est une boule $x^2 + y^2 + z^2 \leqslant R^2$ ou une partie de cette boule.

Exemple 4. Calculer l'intégrale $\iiint_V (x^2+y^2+z^2)\,dx\,dy\,dz$ où V est la boule $x^2 + y^2 + z^2 \leqslant R^2$ (fig. 200).

S o l u t i o n. On a intérêt à passer aux coordonnées sphériques: $x = \rho \sin\theta \cos\varphi$, $y = \rho \sin\theta \sin\varphi$, $z = \rho \cos\theta$. A la forme de V on voit que ρ varie entre 0 et R, φ entre 0 et 2π, θ entre 0 et π. Puisque

$$x^2 + y^2 + z^2 = \rho^2 \sin^2\theta \cos^2\varphi + \rho^2 \sin^2\theta \sin^2\varphi + \rho^2 \cos^2\theta = \rho^2 \sin^2\theta + \rho^2 \cos^2\theta = \rho^2,$$

la formule (5) nous donne

$$\iiint_V (x^2+y^2+z^2)\,dx\,dy\,dz = \int_0^R d\rho \int_0^\pi d\theta \int_0^{2\pi} \rho^2\rho^2 \sin\theta\,d\varphi =$$

$$= \int_0^R \rho^4\,d\rho \int_0^\pi \sin\theta\,d\theta \int_0^{2\pi} d\varphi = 2\pi \int_0^R \rho^4\,d\rho \int_0^\pi \sin\theta\,d\theta =$$

$$= 4\pi \int_0^R \rho^4\,d\rho = 4\pi \frac{\rho^5}{5}\Big|_0^R = \frac{4\pi R^5}{5}.$$

4. Quelques applications des intégrales triples. Voyons brièvement quelques problèmes classiques d'application des intégrales triples nous contentant de citer les formules sans les démontrer en raison de l'analogie de leurs démonstrations avec celles des intégrales doubles.

Si est donné un solide V dont la densité $\rho(M) = \rho(x, y, z)$ est une fonction continue, l'intégrale triple $\iiint_V \rho(x, y, z)\,dx\,dy\,dz$ représente la masse m de V.

Les moments d'inertie d'un solide V, de densité $\rho(M) = \rho(x, y, z)$, par rapport aux axes de coordonnées sont définis par les formules suivantes:

$$I_z = \iiint_V (x^2+y^2)\,\rho(M)\,dv;$$

$$I_y = \iiint_V (x^2+z^2)\,\rho(M)\,dv;$$

$$I_x = \iiint_V (y^2+z^2)\,\rho(M)\,dv.$$

Le moment d'inertie par rapport à l'origine des coordonnées est

$$I_0 = \iiint_V (x^2+y^2+z^2)\,\rho(M)\,dv.$$

Les coordonnées du centre de gravité sont données par les formules

$$x_c = \frac{\iiint_V x\rho(M)\,dv}{m}; \quad y_c = \frac{\iiint_V y\rho(M)\,dv}{m}; \quad z_c = \frac{\iiint_V z\rho(M)\,dv}{m},$$

où m est la masse de V. En particulier, si le solide envisagé est homogène, c'est-à-dire si $\rho(x, y, z) = \text{const}$, les expressions des coordonnées du centre de gravité prennent la forme plus simple

$$x_c = \frac{\iiint_V x\,dv}{v}; \quad y_c = \frac{\iiint_V y\,dv}{v}; \quad z_c = \frac{\iiint_V z\,dv}{v},$$

où v est le volume de V.

Comme déjà signalé, l'intégrale triple $\iiint_V dx\,dy\,dz$ est égale au volume de V. Dans certains cas les intégrales triples se prêtent mieux au calcul des volumes que les intégrales doubles, car elles permettent de calculer le volume aussi bien d'un cylindre curviligne que d'autres corps.

Exemple 5. Déterminer les coordonnées du centre de gravité del'hémisphère boréal V d'une boule homogène de rayon R centrée en l'origine des coordonnées.

Solution. Cet hémisphère est limité par les surfaces $z = \sqrt{R^2 - x^2 - y^2}$ et $z = 0$. Pour cause de symétrie de l'hémisphère, $x_c = y_c = 0$. La coordonnée z_c est donnée par la formule

$$z_c = \frac{\iiint_V z\,dx\,dy\,dz}{\iiint_V dx\,dy\,dz} = \frac{\iiint_V z\,dx\,dy\,dz}{\frac{2}{3}\pi R^3}.$$

En passant aux coordonnées sphériques, on trouve

$$z_c = \frac{\int_0^{2\pi} d\varphi \int_0^{\pi/2} \sin\theta\cos\theta\,d\theta \int_0^R \rho^3\,d\rho}{\frac{2}{3}\pi R^3} = \frac{2\pi\,\frac{R^4}{4}\cdot\frac{1}{2}}{\frac{2}{3}\pi R^3} = \frac{3}{8}R.$$

§ 11. Intégrales de surface

Dans ce paragraphe on étudie les intégrales de fonctions définies sur des surfaces que l'on appelle *intégrales de surface*.

La théorie des intégrales de surface présente de nombreuses affinités avec celle des intégrales curvilignes. On distingue des intégrales de surface de première et de seconde espèce.

1. Définition de l'intégrale de surface de première espèce. Soit $f(M) = f(x, y, z)$ une fonction bornée définie sur une surface S différentiable ou différentiable par morceaux *). Partageons la surface S en n parties d'aires $\Delta S_1, \Delta S_2, \ldots, \Delta S_n$ (fig. 201). Choisissons un point arbitraire $M_i(\xi_i; \eta_i; \zeta_i)$ dans chaque surface partielle et formons la somme

$$\sum_{i=1}^{n} f(M_i)\,\Delta S_i. \qquad (1)$$

La somme (1) s'appelle *somme intégrale* pour la fonction $f(M)$ étendue à la surface S. Désignons par λ le plus grand des diamètres des surfaces partielles.

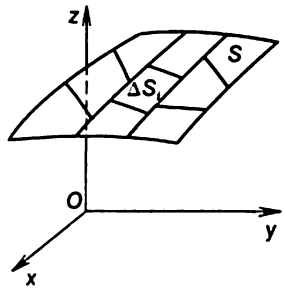


Fig. 201

Définition. La limite I, si elle existe, de la somme intégrale (1) lorsque $\lambda \to 0$ s'appelle *intégrale de surface de première espèce* de la fonction $f(x, y, z)$ étendue à la surface S et se représente par l'une des notations suivantes:

$$I = \iint_S f(M)\,dS = \iint_S f(x, y, z)\,dS.$$

Cette définition est identique à celle de l'intégrale double. C'est pourquoi les propriétés des intégrales doubles et les conditions de leur existence se généralisent pratiquement sans changements aux intégrales de surface.

En particulier, si $f(x, y, z) \equiv 1$ sur la surface S, alors

$$\iint_S dS = \lim_{\lambda \to 0} \sum_{i=1}^{n} \Delta S_i = s,$$

où s est l'aire de S, autrement dit l'intégrale de surface de première espèce peut servir au calcul des aires de surface.

D'autre part, on peut les utiliser pour calculer les masses, les moments statiques, les moments d'inertie, les coordonnées du centre de gravité des surfaces matérielles dont la densité surfacique est connue. Ces problèmes se résolvent comme pour une courbe matérielle, un domaine matériel plan et gauche.

---

*) On dit qu'une surface est *différentiable* si elle admet un plan tangent en chacun de ses points et si ce plan varie continûment quand on passe d'un point à un autre. Une surface composée d'un nombre fini de morceaux différentiables s'appelle *différentiable par morceaux*.

2. Calcul des intégrales de surface de première espèce. L'intégrale de surface de première espèce se calcule par réduction à l'intégrale double.

Soit S une surface représentative d'une fonction $z = z(x, y)$ continue avec ses dérivées $z'_x(x, y)$ et $z'_y(x, y)$ sur le domaine fermé G, projeté de S sur le plan Oxy (fig. 202); et soit $f(x, y, z)$ une fonction continue sur S, donc intégrable sur S.

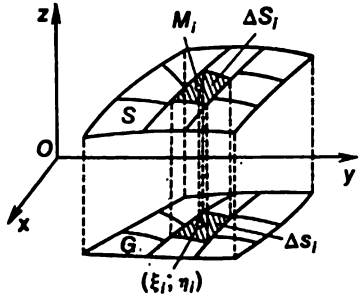


Fig. 202

Partageons la surface S arbitrairement en n parties et projetons cette subdivision sur le plan Oxy. Nous obtenons une subdivision du domaine G en domaines partiels $G_1, G_2, \ldots, G_n$. L'aire ΔS_i de chaque surface partielle s'exprime par la formule (cf. formule (2), n° 3, § 4)

$$\Delta S_i = \iint_{G_i} \sqrt{1 + z'^2_x(x, y) + z'^2_y(x, y)}\, dx\, dy.$$

En appliquant le théorème de la moyenne à cette intégrale double, on obtient

$$\Delta S_i = \sqrt{1 + z'^2_x(\xi_i, \eta_i) + z'^2_y(\xi_i, \eta_i)}\, \Delta s_i, \qquad (2)$$

où $(\xi_i; \eta_i)$ est un point du domaine G_i; Δs_i, l'aire de G_i. Désignons par M_i le point de la surface S_i de coordonnées $(\xi_i; \eta_i; \zeta_i)$, où $\zeta_i = z(\xi_i, \eta_i)$ et $(\xi_i; \eta_i)$ est le point intervenant dans la formule (2). Formons la somme intégrale pour la fonction $f(x, y, z)$ sur la surface S en choisissant les points M_i pour points intermédiaires :

$$\sum_{i=1}^{n} f(\xi_i, \eta_i, \zeta_i)\, \Delta S_i =$$

$$= \sum_{i=1}^{n} f[\xi_i, \eta_i, z(\xi_i, \eta_i)] \sqrt{1 + z'^2_x(\xi_i, \eta_i) + z'^2_y(\xi_i, \eta_i)}\, \Delta s_i. \qquad (3)$$

Le second membre est la somme intégrale de l'intégrale double de la fonction $f[x, y, z(x, y)] \sqrt{1 + z'^2_x(x, y) + z'^2_y(x, y)}$ continue sur le domaine G. Donc la limite du second membre de (3) lorsque $\lambda \to 0$ est égale à l'intégrale double

$$\iint_G f[x, y, z(x, y)] \sqrt{1 + z'^2_x(x, y) + z'^2_y(x, y)}\, dx\, dy.$$

La fonction $f(x, y, z)$ étant intégrable sur la surface S, la limite du premier membre de (3) lorsque $\lambda \to 0$ est égale à l'intégrale de surface

$$\iint_S f(x, y, z)\, dS.$$

Donc en passant à la limite dans (3) ($\lambda \to 0$), on obtient la formule cherchée

$$\iint_S f(x, y, z)\, dS =$$

$$= \iint_G f[x, y, z(x, y)] \sqrt{1 + z'^2_x(x, y) + z'^2_y(x, y)}\, dx\, dy, \quad (4)$$

qui exprime une intégrale de surface de première espèce par l'intermédiaire d'une intégrale double étendue à la projetée de la surface S sur le plan Oxy.

On obtient de façon analogue les formules exprimant l'intégrale étendue à la surface S par l'intermédiaire d'intégrales doubles étendues aux projetées de S sur les plans Oyz et Oxz.

Exemple 1. Calculer l'intégrale $\iint_S \sqrt{1 + 4x^2 + 4y^2}\, dS$, où S est la portion de paraboloïde de révolution $z = 1 - x^2 - y^2$ découpée par le plan $z = 0$ (fig. 203).

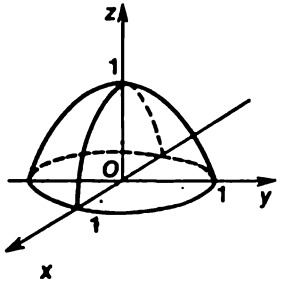


Fig. 203

S o l u t i o n. La projetée de la surface S d'équation $z = 1 - x^2 - y^2$ sur le plan Oxy est le cercle $x^2 + y^2 = 1$ (c'est-à-dire la section $z = 0$ du paraboloïde). Donc le domaine G est le disque $x^2 + y^2 \leqslant 1$. La fonction $z = 1 - x^2 - y^2$ et ses dérivées $z'_x(x, y) = -2x$ et $z'_y(x, y) = -2y$ étant continues sur G, la formule (4) nous donne

$$\iint_S f(x, y, z)\, dS = \iint_S \sqrt{1 + 4x^2 + 4y^2}\, dS =$$

$$= \iint_G \sqrt{1 + 4x^2 + 4y^2} \sqrt{1 + 4x^2 + 4y^2}\, dx\, dy =$$

$$= \iint_G (1 + 4x^2 + 4y^2)\, dx\, dy.$$

En passant aux coordonnées polaires $x = \rho \cos \varphi$, $y = \rho \sin \varphi$, on trouve

$$\iint_G (1 + 4x^2 + 4y^2)\, dx\, dy = \int_0^{2\pi} d\varphi \int_0^1 (1 + 4\rho^2)\, \rho\, d\rho =$$

$$= \int_0^{2\pi} \left[\frac{\rho^2}{2} + \rho^4\right]_0^1 d\varphi = \frac{3}{2} \int_0^{2\pi} d\varphi = \frac{3}{2}\, \varphi \Big|_0^{2\pi} = 3\pi.$$

3. Définition de l'intégrale de surface de seconde espèce. Introduisons préalablement la notion de côté d'une surface.

Soient S une surface différentiable, M un point arbitraire de S, $\vec{n}$ le vecteur normal à S en M. Considérons maintenant un contour fermé sur S passant par le point M et ne rencontrant pas la

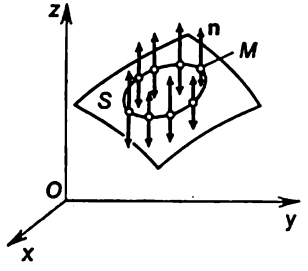


Fig. 204

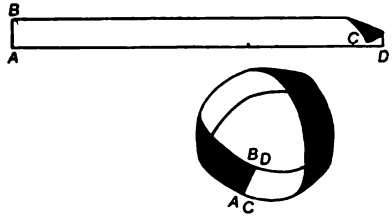


Fig. 205

frontière de S. Déplaçons l'origine M du vecteur $\vec{n}$ le long du contour fermé de telle sorte que $\vec{n}$ reste normal à S et que sa direction varie continûment *) (fig. 204). Le vecteur $\vec{n}$ revient à sa position de départ avec son sens initial ou le sens opposé.

On dit qu'une surface est à deux côtés si, en se déplaçant le long de tout contour fermé, le vecteur normal revient à sa position de départ en conservant son sens initial.

Comme exemples de surfaces à deux côtés citons le plan, la sphère, toute surface d'équation $z = f(x, y)$, où $f(x, y)$ est une fonction continue avec ses dérivées premières sur un domaine du plan Oxy.

On dit qu'une surface S est à un côté s'il existe dans S un contour fermé le long duquel le vecteur normal change de sens en revenant à sa position de départ.

Un exemple élémentaire de surface à un seul côté est le ruban de Möbius **) qui est représenté sur la figure 205. On l'obtient

*) Cela signifie que les cosinus directeurs de $\vec{n}$ sont des fonctions continues des coordonnées de S.

**) Möbius August Ferdinand (1790-1868), mathématicien allemand.

en tordant une fois une bande de papier rectangulaire $ABCD$ et en collant le côté DC sur le côté BA. Lorsque le vecteur normal parcourt le ruban le long de la ligne médiane, il revient à sa position de départ avec un sens opposé au sens initial.

Dans la suite on n'étudiera que les surfaces à deux côtés. On appelle *côté d'une surface* S l'ensemble constitué de tous les points de S et d'un vecteur normal à S dont la direction varie continûment d'un point à un autre de S. Le choix d'un côté de S s'appelle *orientation* de S. Les surfaces à deux côtés sont dites *orientables*, les surfaces à un côté, *non orientables*.

A la notion de côté d'une surface est étroitement liée la notion d'orientation du contour de cette surface.

Soit S une surface orientée limitée par un contour L sans point double. On conviendra que le contour L est orienté dans le sens *direct*, ou *positif*, si un observateur se déplaçant sur L de telle sorte que le vecteur normal lui entre par les pieds et sorte par la tête, laisse l'intérieur de L sur sa gauche (fig. 206). Le sens de parcours opposé est dit *rétrograde*, ou *négatif*.

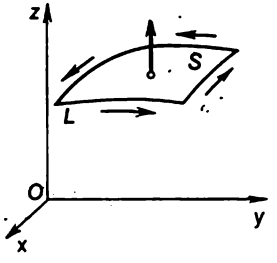


Fig. 206

Passons maintenant à la définition de l'intégrale de surface de seconde espèce.

Soit S une surface différentiable d'équation $z = f(x, y)$ et soit $R(x, y, z)$ une fonction bornée définie sur S. Considérons l'un des côtés de cette surface. On dira que le côté choisi est le *côté supérieur* ou le *côté inférieur* de la surface $z = f(x, y)$ selon que les normales à S font des angles aigus ou obtus avec l'axe Oz. Partageons la surface S arbitrairement en n parties et désignons par G_i la projetée de la surface partielle S_i sur le plan Oxy. Prenons un point arbitraire $M_i(\xi_i; \eta_i; \zeta_i)$ dans chaque surface partielle et composons la somme

$$\sum_{i=1}^{n} R(\xi_i; \eta_i; \zeta_i)\,\Delta s_i, \tag{5}$$

où Δs_i est l'aire de G_i prise avec le signe « + » ou « — » selon que l'on considère le côté supérieur ou le côté inférieur de S. La somme (5) s'appelle *somme intégrale* pour la fonction $R(M) = R(x, y, z)$. Désignons par λ le plus grand des diamètres des surfaces partielles.

Définition. La limite I, si elle existe, de la somme intégrale (5) lorsque $\lambda \to 0$ s'appelle *intégrale de surface de seconde espèce* de la fonction $R(x, y, z)$ étendue au côté choisi de la surface S et se

représente par l'une des notations suivantes:

$$I = \iint_S R(M)\,dx\,dy = \iint_S R(x, y, z)\,dx\,dy.$$

La fonction $R(x, y, z)$ est dite *intégrable sur la surface* S par rapport aux variables x et y.

On définit de façon analogue l'intégrale de surface de seconde espèce étendue au côté choisi de S par rapport aux variables y et z (resp. z et x) d'une fonction $P(x, y, z)$ (resp. $Q(x, y, z)$) définie sur S:

$$\iint_S P(x, y, z)\,dy\,dz \left[\text{resp. } \iint_S Q(x, y, z)\,dz\,dx\right].$$

La somme

$$\iint_S P(x, y, z)\,dy\,dz + \iint_S Q(x, y, z)\,dz\,dx + \iint_S R(x, y, z)\,dx\,dy$$

s'appelle *intégrale de surface générale de seconde espèce* et s'écrit

$$\iint_S P(x, y, z)\,dy\,dz + Q(x, y, z)\,dz\,dx + R(x, y, z)\,dx\,dy. \quad (6)$$

L'intégrale de surface de seconde espèce possède les mêmes propriétés que l'intégrale de surface de première espèce à une différence près: c'est qu'elle change de signe quand on passe d'un côté à l'autre de la surface.

Le problème du flux d'un champ de vecteurs qui sera envisagé au § 14 conduit à la notion d'intégrale de surface de seconde espèce.

Cette notion n'est pas introduite pour les surfaces à un côté.

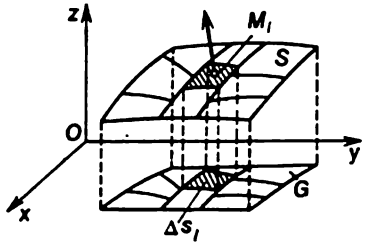


Fig. 207

4. Calcul des intégrales de surface de seconde espèce. Les intégrales de surface de seconde espèce se calculent par réduction à des intégrales doubles.

Soit S une surface différentiable orientée par le choix du côté supérieur, représentative d'une fonction $z = f(x, y)$ définie sur un domaine fermé G, projeté de S sur le plan Oxy, et soit $R(x, y, z)$ une fonction continue sur S.

Partageons la surface S arbitrairement en n parties et projetons cette subdivision sur le plan Oxy (fig. 207). Le domaine se subdivise alors en n parties $G_1, G_2, \ldots, G_n$. Choisissons dans chaque

surface partielle un point arbitraire $M_i(\xi_i; \eta_i; \zeta_i)$ et formons la somme intégrale $\sum_{i=1}^{n} R(\xi_i, \eta_i, \zeta_i)\,\Delta s_i$, où Δs_i est l'aire de G_i. Puisque $\zeta_i = f(\xi_i, \eta_i)$, on a

$$\sum_{i=1}^{n} R(\xi_i, \eta_i, \zeta_i)\,\Delta s_i = \sum_{i=1}^{n} R[\xi_i, \eta_i, f(\xi_i, \eta_i)]\,\Delta s_i. \qquad (7)$$

Le second membre est la somme intégrale de l'intégrale double de la fonction $R[x, y, f(x, y)]$ continue sur le domaine G. En faisant tendre λ vers 0 dans (7) on obtient la formule cherchée

$$\iint_S R(x, y, z)\,dx\,dy = \iint_G R[x, y, f(x, y)]\,dx\,dy, \qquad (8)$$

qui exprime l'intégrale de surface de seconde espèce par rapport aux variables x et y au moyen d'une intégrale double. Par ailleurs, la formule (8) établit l'existence de l'intégrale de surface de la fonction $R(x, y, z)$ continue sur la surface S envisagée. Si l'on se place sur le côté inférieur de S, l'intégrale du second membre de (8) change de signe.

On établit de façon analogue la validité des formules suivantes :

$$\iint_S P(x, y, z)\,dy\,dz = \iint_{G_1} P[f(y, z), y, z]\,dy\,dz, \qquad (9)$$

$$\iint_S Q(x, y, z)\,dz\,dx = \iint_{G_2} Q[x, f(x, z), z]\,dz\,dx, \qquad (10)$$

où la surface S est définie respectivement par les équations $x = f(x, y)$ et $y = f(x, z)$ et G_1 et G_2 sont les projetées de S respectivement sur les plans Oyz et Oxz.

Pour calculer l'intégrale générale (6) on se sert des mêmes formules (8), (9) et (10) si la surface se projette de façon unique sur les trois plans de coordonnées. Dans les cas plus compliqués la surface S est partagée en parties possédant les propriétés indiquées, et l'intégrale (6), en une somme d'intégrales étendues à ces parties.

Exemple 2. Calculer l'intégrale $\iint_S (y^2 + z^2)\,dx\,dy$, où S est le côté supérieur de la surface $z = \sqrt{1 - x^2}$ limitée par les plans $y = 0$ et $y = 1$ (fig. 208).

Solution. La projetée G de cette surface sur le plan Oxy est le rectangle $-1 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$. La formule (8) nous

donne

$$\iint_S (y^2+z^2)\,dx\,dy = \iint_G [y^2+(\sqrt{1-x^2})^2]\,dx\,dy =$$

$$= \int_{-1}^{1} dx \int_0^1 (y^2+1-x^2)\,dy = \int_{-1}^{1} \left[\frac{y^3}{3}+y-x^2y\right]_0^1 dx =$$

$$= \int_{-1}^{1} \left(\frac{4}{3}-x^2\right) dx = \left[\frac{4}{3}x-\frac{x^3}{3}\right]_{-1}^{1} = 2.$$

Exemple 3. Calculer l'intégrale $\iint_S x\,dy\,dz + y\,dz\,dx + z\,dx\,dy$, où S est le côté supérieur de la région du plan $x + z - 1 = 0$ li-

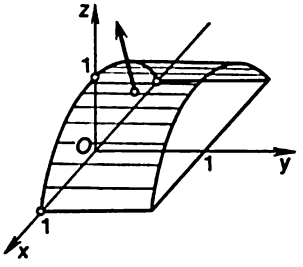


Fig. 208

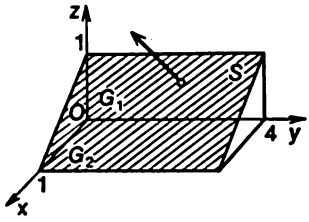


Fig. 209

mitée par les plans $y = 0$ et $y = 4$, et située dans le premier octant (fig. 209).

Solution. Par définition,

$$\iint_S x\,dy\,dz + y\,dz\,dx + z\,dx\,dy =$$

$$= \iint_{G_1} x(y,\ z)\,dy\,dz + \iint_S y\,dz\,dx + \iint_{G_2} z(x,\ y)\,dx\,dy,$$

où G_1 et G_2 sont les projetées de S sur les plans Oyz et Oxy et $\iint_S y\,dz\,dx = 0$, puisque la surface S est parallèle à Oy. Les formules (8) et (9) nous donnent respectivement

$$\iint_S z\,dx\,dy = \iint_{G_2} (1-x)\,dx\,dy = \int_0^4 dy \int_0^1 (1-x)\,dx = 2,$$

$$\iint_S x\,dy\,dz = \iint_{G_1} (1-z)\,dy\,dz = \int_0^4 dy \int_0^1 (1-z)\,dz = 2.$$

Donc

$$\iint_S x\,dy\,dz + y\,dz\,dx + z\,dx\,dy = 2 + 0 + 2 = 4.$$

5. Lien entre les intégrales de surface de première et de seconde espèce. Les intégrales de surface de seconde espèce peuvent être introduites d'une autre manière, plus exactement, comme des intégrales de surface de première espèce contenant des expressions spéciales sous le signe d'intégration. Désignons par $\cos\alpha$, $\cos\beta$ et $\cos\gamma$ les cosinus directeurs de la normale à une surface en l'un quelconque de ses points. Les intégrales de surface de seconde espèce diffèrent entre elles par le plan sur lequel est projetée la surface S :

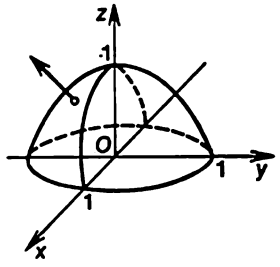


Fig. 210

1) pour le plan Oxy

$$\iint_S R(x, y, z)\,dx\,dy = \iint_S R(x, y, z)\cos\gamma\,dS; \qquad (11)$$

2) pour le plan Oxz

$$\iint_S Q(x, y, z)\,dz\,dx = \iint_S Q(x, y, z)\cos\beta\,dS; \qquad (12)$$

3) pour le plan Oyz

$$\iint_S P(x, y, z)\,dy\,dz = \iint_S P(x, y, z)\cos\alpha\,dS. \qquad (13)$$

En sommant les formules (11), (12) et (13) on obtient une formule exprimant l'intégrale de surface générale de seconde espèce étendue au côté choisi de la surface, au moyen d'une intégrale de surface de première espèce :

$$\iint_S P\,dy\,dz + Q\,dz\,dx + R\,dx\,dy =$$

$$= \iint_S (P\cos\alpha + Q\cos\beta + R\cos\gamma)\,dS. \qquad (14)$$

Si l'on se place sur l'autre côté de la surface, les cosinus directeurs $\cos\alpha$, $\cos\beta$ et $\cos\gamma$ de la normale changent de signe, donc l'intégrale de surface de seconde espèce aussi.

Exemple 4. Calculer l'intégrale $\iint_S z\cos\gamma\,dS$, où S est le côté extérieur de l'hémisphère boréal $x^2 + y^2 + z^2 = 1$. γ l'angle aigu de la normale à S et de l'axe Oz (fig. 210).

Solution. La formule (11) qui relie les intégrales de surface des deux espèces nous donne

$$\iint_S z \cos\gamma \, dS = \iint_S z \, dx \, dy.$$

La projetée G de la surface S sur le plan Oxy est le disque $x^2 + y^2 \leqslant 1$. Donc, d'après la formule (8),

$$\iint_S z \, dx \, dy = \iint_G \sqrt{1 - x^2 - y^2} \, dx \, dy.$$

En passant aux coordonnées polaires dans l'intégrale double, on obtient

$$\iint_G \sqrt{1 - x^2 - y^2} \, dx \, dy = \int_0^{2\pi} d\varphi \int_0^1 \sqrt{1 - \rho^2} \, \rho \, d\rho =$$

$$= \int_0^{2\pi} \left[-\frac{(1 - \rho^2)^{3/2}}{3} \right]_0^1 d\varphi = \frac{1}{3} \int_0^{2\pi} d\varphi = \frac{2}{3} \pi.$$

§ 12. Formule d'Ostrogradski

La formule d'Ostrogradski *) établit une relation entre une intégrale de surface étendue à une surface fermée et l'intégrale triple étendue au volume limité par cette surface. Cette formule est analogue à la formule de Green qui, rappelons-le, relie une intégrale curviligne le long d'une courbe fermée à une intégrale double étendue au domaine plan limité par cette courbe. La formule d'Ostrogradski joue un rôle important en analyse et dans ses applications.

Etablissons cette formule pour un domaine gauche fermé dont la frontière est coupée en deux points au plus par toute droite parallèle aux axes de coordonnées. Pour abréger, on dira que ces domaines sont simples. On se placera sur le côté extérieur de la surface limitant ce domaine et on admettra que cette surface est différentiable ou différentiable par morceaux.

Théorème 13.8. *Soit V un volume fermé simple limité par une surface S et soient $P(x, y, z)$, $Q(x, y, z)$ et $R(x, y, z)$ des fonctions continues avec leurs dérivées partielles premières sur V. On a alors*

*) Ostrogradski Mikhaïl (1801-1861), mathématicien russe

la formule suivante:

$$\iiint_V \left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right) dx\,dy\,dz = \iint_S P\,dy\,dz+Q\,dz\,dx+R\,dx\,dy, \quad (1)$$

dite formule d'Ostrogradski.

Démonstration. Soient G la projetée de la surface S (et du volume V) sur le plan Oxy (fig. 211), $z = z_1(x, y)$ et $z = z_2(x, y)$, les équations des parties inférieure S_1 et supérieure S_2 de la surface S.

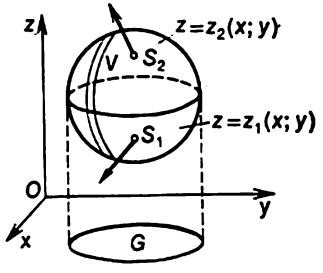


Fig. 211

Transformons l'intégrale triple

$$\iiint_V \frac{\partial R}{\partial z}\,dx\,dy\,dz$$

en intégrale de surface. A cet effet, ramenons-la à une intégrale itérée et intégrons par rapport à z avec la formule de Newton-Leibniz. On obtient

$$\iiint_V \frac{\partial R}{\partial z}\,dx\,dy\,dz = \iint_G dx\,dy \int_{z_1(x,\,y)}^{z_2(x,\,y)} \frac{\partial R}{\partial z}\,dz =$$

$$= \iint_G R[x, y, z_2(x, y)]\,dx\,dy - \iint_G R[x, y, z_1(x, y)]\,dx\,dy.$$

Le domaine G étant le projeté sur le plan Oxy et de la surface S_2, et de la surface S_1, les intégrales doubles peuvent être remplacées par des intégrales de surfaces égales [cf. § 11, n° 4 formule (8)], étendues aux côtés extérieurs des surfaces $z = z_2(x, y)$ et $z = z_1(x, y)$, c'est-à-dire

$$\iiint_V \frac{\partial R}{\partial z}\,dx\,dy\,dz = \iint_{S_2} R(x, y, z)\,dx\,dy -$$

$$- \iint_{S_1} R(x, y, z)\,dx\,dy.$$

Dans l'intégrale étendue à S_1, en changeant de côté on obtient

$$\iiint_V \frac{\partial R}{\partial z}\,dx\,dy\,dz = \iint_{S_2} R(x, y, z)\,dx\,dy +$$

$$+ \iint_{S_1} R(x, y, z)\,dx\,dy = \iint_S R\,dx\,dy, \quad (2)$$

où S est le côté extérieur de la surface limitant le volume V.

On démontre de façon analogue les formules

$$\iiint_V \frac{\partial P}{\partial x}\,dx\,dy\,dz = \iint_S P\,dy\,dz, \tag{3}$$

$$\iiint_V \frac{\partial Q}{\partial y}\,dx\,dy\,dz = \iint_S Q\,dz\,dx. \tag{4}$$

En ajoutant terme à terme les égalités (2), (3) et (4), on obtient la formule (1). ■

Remarque. La formule d'Ostrogradski est valable pour tout volume fermé V divisible en un nombre fini de volumes partiels. En effet, en appliquant la formule (1) à chaque volume partiel de la subdivision et en ajoutant les résultats, on obtient au premier membre une intégrale triple étendue à V et au second, une intégrale de surface étendue à la surface S limitant V, puisque la somme des intégrales de surface étendues aux deux côtés des surfaces auxiliaires est nulle.

La formule d'Ostrogradski se prête bien au calcul des intégrales de surface étendues à des surfaces fermées.

Exemple 1. Calculer l'intégrale $\iint_S x\,dy\,dz + y\,dz\,dx + z\,dx\,dy$, où S est le côté extérieur de la pyramide limitée par les plans $x + y + z = 1$, $x = 0$, $y = 0$ et $z = 0$ (cf. fig. 196).

Solution. La formule d'Ostrogradski nous donne

$$\iint_S x\,dy\,dz + y\,dz\,dx + z\,dx\,dy = \iiint_V (1+1+1)\,dx\,dy\,dz =$$

$$= 3\iiint_V dx\,dy\,dz = 3\int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz =$$

$$= 3\int_0^1 dx \int_0^{1-x} [z]_0^{1-x-y}\,dy = 3\int_0^1 \left[y - xy - \frac{y^2}{2}\right]_0^{1-x} dx =$$

$$= 3\int_0^1 \left[1 - x - x(1-x) - \frac{(1-x)^2}{2}\right] dx = 3 \cdot \frac{1}{6} = \frac{1}{2}.$$

Exemple 2. Calculer l'intégrale $\iint_S x^3\,dy\,dz + y^3\,dz\,dx + z^3\,dx\,dy$, où S est le côté extérieur de la sphère $x^2 + y^2 + z^2 = R^2$.

Solution. La formule d'Ostrogradski nous donne

$$\iint_S x^3\,dy\,dz + y^3\,dz\,dx + z^3\,dx\,dy = 3\iiint_V (x^2 + y^2 + z^2)\,dx\,dy\,dz,$$

d'où, si l'on passe aux coordonnées sphériques,

$$3\iiint_V (x^2+y^2+z^2)\,dx\,dy\,dz =$$

$$= 3\int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta\,d\theta \int_0^{R} \rho^4\,d\rho = \frac{12}{5}\pi R^5.$$

Comme déjà signalé (§ 9, n° 1), la formule de Green exprime l'aire d'un domaine au moyen d'une intégrale curviligne le long du contour de ce domaine. De façon exactement analogue on peut déduire de la formule d'Ostrogradski une expression pour le volume sous forme d'une intégrale de surface étendue à la surface fermée S limitant ce volume. En effet, en choisissant des fonctions P, Q et R telles que

$$\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\equiv 1,$$

on obtient

$$v=\iiint_V dx\,dy\,dz = \iint_S P\,dy\,dz+Q\,dz\,dx+R\,dx\,dy,$$

où v est le volume limité par S. En particulier, en posant $P = x/3$, $Q = y/3$ et $R = z/3$, on a

$$v=\frac{1}{3}\iint_S x\,dy\,dz+y\,dz\,dx+z\,dx\,dy.$$

§ 13. Formule de Stokes

La formule de Stokes *) établit un lien entre les intégrales de surface et les intégrales curvilignes. Comme les formules de Green et d'Ostrogradski, la formule de Stokes est largement utilisée en analyse et dans ses applications.

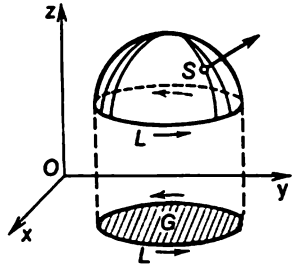


Fig. 212

Soient S une surface représentative d'une fonction $z = z(x, y)$, continue avec ses dérivées partielles premières $z'_x(x, y)$ et $z'_y(x, y)$ sur le domaine fermé G, projeté de S sur le plan Oxy; L le contour limitant S; l le contour de G. Plaçons-nous sur le côté supérieur de la surface S (fig. 212). Sous ces conditions on a le

Théorème 13.9. *Si une fonction $P(x, y, z)$ est continue avec ses dérivées partielles premières sur la surface S, on a la formule sui-*

*) Stokes Georges Gabriel (1819-1903), mathématicien et physicien anglais.

vante :

$$\oint_L P(x, y, z)\,dx = \iint_S \left(\frac{\partial P}{\partial z}\cos\beta - \frac{\partial P}{\partial y}\cos\gamma\right) dS, \qquad (1)$$

où $\cos\beta$ *et* $\cos\gamma$ *sont les cosinus directeurs de la normale à* S, *le contour* L *est parcouru dans le sens direct.*

Démonstration. Transformons l'intégrale curviligne

$$\oint_L P(x, y, z)\,dx$$

en intégrale étendue à la surface S. Cette transformation s'effectue d'après le schéma suivant :

$$\oint_L \to \oint_l \to \iint_G \to \iint_S .$$

Le contour L étant situé sur la surface S, ses points satisfont à l'équation $z = z(x, y)$ et par suite la fonction $P(x, y, z)$ prend aux points de L des valeurs égales à celles de la fonction $P[x, y, z(x, y)]$ aux points correspondants de l. Les projetées des subdivisions de L et de l sur l'axe Ox sont confondues. Donc les sommes intégrales des intégrales curvilignes de seconde espèce de la fonction P le long des contours L et l sont aussi confondues et par suite

$$\oint_L P(x, y, z)\,dx = \oint_l P[x, y, z(x, y)]\,dx.$$

En appliquant la formule de Green, on obtient

$$\oint_l P[x, y, z(x, y)]\,dx = -\iint_G \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\,z'_y\right) dx\,dy.$$

L'intégrant est ici la dérivée partielle par rapport à y de la fonction composée $P[x, y, z(x, y)]$.

Puisque l'on considère le côté supérieur de la surface S ($\cos\gamma > 0$), les composantes de la normale sont $\{-z'_x, -z'_y, 1\}$. Les cosinus directeurs de la normale étant proportionnels à ces composantes, on a

$$\frac{\cos\beta}{\cos\gamma} = \frac{-z'_y}{1} = -z'_y.$$

Donc

$$-\iint_G \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\,z'_y\right) dx\,dy = -\iint_G \left(\frac{\partial P}{\partial y} - \frac{\partial P}{\partial z}\,\frac{\cos\beta}{\cos\gamma}\right) dx\,dy.$$

Les formules (8) et (11) du § 11 nous permettent maintenant de transformer cette intégrale double en intégrale de surface :

$$-\iint_G \left(\frac{\partial P}{\partial y}-\frac{\partial P}{\partial z}\,\frac{\cos\beta}{\cos\gamma}\right) dx\,dy=-\iint_S \left(\frac{\partial P}{\partial y}\cos\gamma-\frac{\partial P}{\partial z}\cos\beta\right) dS.$$

En définitive,

$$\oint_L P(x,\ y,\ z)\,dx=\iint_S \left(\frac{\partial P}{\partial z}\cos\beta-\frac{\partial P}{\partial y}\cos\gamma\right) dS. \blacksquare$$

On démontre de façon analogue *mutatis mutandis* les trois formules suivantes :

$$\oint_L Q(x,\ y,\ z)\,dy=\iint_S \left(\frac{\partial Q}{\partial x}\cos\gamma-\frac{\partial Q}{\partial z}\cos\alpha\right) dS. \tag{2}$$

$$\oint_L R(x,\ y,\ z)\,dz=\iint_S \left(\frac{\partial R}{\partial y}\cos\alpha-\frac{\partial R}{\partial x}\cos\beta\right) dS. \tag{3}$$

En ajoutant membre à membre les égalités (1), (2) et (3), on obtient la formule

$$\oint_L P\,dx+Q\,dy+R\,dz=\iint_S \left[\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\cos\gamma+\right.$$
$$\left.+\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\cos\alpha+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\cos\beta\right] dS,$$

dite *formule de Stokes.*

Grâce à la formule (14) du § 11, la formule de Stokes peut encore s'écrire

$$\oint_L P\,dx+Q\,dy+R\,dz=\iint_S \left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right) dx\,dy+$$
$$+\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right) dy\,dz+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right) dz\,dx. \tag{4}$$

On peut facilement retenir la formule de Stokes en remarquant que le premier terme du second membre n'est autre que l'intégrale double de la formule de Green et que le deuxième et le troisième s'en déduisent par une permutation circulaire des variables x, y, z et des fonctions P, Q, R.

En particulier, si la surface S est un domaine du plan Oxy limité par un contour L, les intégrales par rapport à $dz\,dx$ et $dy\,dz$ sont nulles et la formule de Stokes se transforme en la formule de Green.

La formule de Stokes permet de calculer des intégrales curvilignes le long de contours fermés au moyen d'intégrales de surface.

Exemple. Calculer à l'aide de la formule de Stokes l'intégrale $\oint\limits_L x^2y^3\,dx + dy + z\,dz$, où L est le cercle $x^2+y^2=1$, $z=0$ et la surface S, l'hémisphère boréal $x^2+y^2+z^2=1$, le contour L étant parcouru dans le sens direct.

Solution. Puisque

$$\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}=-3x^2y^2;\quad \frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}=0;\quad \frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}=0,$$

la formule de Stokes (4) nous donne

$$\oint\limits_L x^2y^3\,dx+dy+z\,dz=-3\iint\limits_S x^2y^2\,dx\,dy=-\frac{\pi}{8}.$$

De la formule de Stokes il s'ensuit que si

$$\frac{\partial Q}{\partial x}=\frac{\partial P}{\partial y};\quad \frac{\partial R}{\partial y}=\frac{\partial Q}{\partial z};\quad \frac{\partial P}{\partial z}=\frac{\partial R}{\partial x}, \tag{5}$$

alors

$$\oint\limits_L P\,dx+Q\,dy+R\,dz=0 \tag{6}$$

quelle que soit la courbe fermée gauche L. Or, ceci exprime que dans ce cas l'intégrale curviligne ne dépend pas du chemin d'intégration.

Comme dans le cas d'une courbe plane, les conditions (5) sont nécessaires et suffisantes pour que l'égalité (6) ait lieu.

Si les conditions (5) ou (6) sont remplies, l'expression $P\,dx+Q\,dy+R\,dz$ est la différentielle totale d'une fonction $U(x, y, z)$:

$$dU=P\,dx+Q\,dy+R\,dz,$$

et par suite

$$\int\limits_{(x_0;\,y_0;\,z_0)}^{(x_1;\,y_1;\,z_1)} P\,dx+Q\,dy+R\,dz=U(x_1;\,y_1;\,z_1)-U(x_0;\,y_0;\,z_0).$$

Cette relation se prouve comme la formule correspondante (4) du § 8 pour les fonctions de deux variables.

§ 14. Champs scalaire et vectoriel

La notion de champ est à la base de nombreux concepts de physique moderne. L'étude de la théorie du champ débordant le cadre de ce cours, on se bornera à de brèves généralités.

D'une façon générale, on dit qu'un champ d'une grandeur u est défini sur un espace si cette grandeur est définie en chaque point de cet espace (ou d'une partie de cet espace). Ainsi, par

exemple, quand on étudie un écoulement de gaz, on doit envisager plusieurs champs: le champ de température (la température prend une valeur définie en chaque point), le champ de pression, le champ de vitesse, etc.

Le champ de u est dit *stationnaire* si u ne dépend pas du temps t, et *non stationnaire* dans le cas contraire. Ainsi la grandeur u est une fonction du point M et du temps t.

Les problèmes de physique font souvent intervenir des grandeurs scalaires et vectorielles. On envisagera donc des *champs scalaires* et des *champs vectoriels* qui seront supposés stationnaires par souci de simplicité.

1. Champ scalaire. Soit G un domaine du plan ou de l'espace. On dit qu'un champ scalaire *est défini* sur le domaine G si une grandeur scalaire u est définie en chaque point M de G. Les notions de champ scalaire et de fonction scalaire définie sur G sont confondues. On définit généralement un champ scalaire par une fonction $u = F(M)$, appelée *fonction scalaire*. Si l'espace est rapporté à un système de coordonnées rectangulaires $Oxyz$, chaque point M est défini par ses coordonnées x, y, z et la grandeur scalaire u sera fonction de ces coordonnées: $u = F(M) = F(x, y, z)$.

Comme exemple de champ scalaire on pourrait citer le champ de la température de l'air dans un local si cette température est traitée comme une fonction de point. Au voisinage de la source de chaleur la température est plus élevée qu'ailleurs. Si la température est partout la même, le champ scalaire est constant.

2. Champ vectoriel. La notion de champ vectoriel s'introduit comme celle d'un champ scalaire: on dit qu'un champ vectoriel est défini sur un domaine G si un vecteur $\vec{F}(M)$ est défini en chaque point M de G. La fonction $\vec{F}(M)$ qui a servi à définir le champ vectoriel s'appelle *fonction vectorielle*.

Un exemple de champ vectoriel nous est fourni par tout champ de force. A chaque point du domaine correspond un vecteur de même direction, de même sens et de même intensité que la force appliquée en ce point.

Exemple 1. Trouver le champ de vitesse $\vec{v}(M)$ des points d'un corps tournant à une vitesse angulaire constante ω autour d'un axe.

Solution. La vitesse $\vec{v}$ d'un point M est égale au produit vectoriel $\vec{v} = \vec{\omega} \wedge \vec{r}$, où $\vec{\omega}$ est le vecteur vitesse angulaire, $\vec{r}$, le rayon vecteur de M par rapport à un point de l'axe de rotation. Prenons ce point pour origine des coordonnées et l'axe de rota-

tion pour axe Oz. Alors $\vec{\omega} = \omega\vec{k}$, $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$ et par suite

$$= \vec{\omega} \wedge \vec{r} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 0 & 0 & \omega \\ x & y & z \end{vmatrix} =$$

$$= \begin{vmatrix} 0 & \omega \\ y & z \end{vmatrix} \vec{i} - \begin{vmatrix} 0 & \omega \\ x & z \end{vmatrix} \vec{j} + \begin{vmatrix} 0 & 0 \\ x & y \end{vmatrix} \vec{k} = -\omega y\vec{i} + \omega x\vec{j}$$

est le champ vectoriel cherché.

3. Champ potentiel. Introduisons la notion de champ potentiel. Considérons un champ scalaire $F(M)$. Si en chaque point M de G est défini le vecteur $\overrightarrow{\text{grad}}\, F$, le champ de ce vecteur s'appelle *champ potentiel*. Le champ scalaire $F(M)$ s'appelle alors *potentiel du champ vectoriel* et le vecteur définissant le champ potentiel, *vecteur potentiel*, autrement dit un vecteur $\vec{a}(M)$ est potentiel s'il existe une fonction scalaire $F(M)$ telle que

$$\vec{a} = \overrightarrow{\text{grad}}\, F = \frac{\partial F}{\partial x}\vec{i} + \frac{\partial F}{\partial y}\vec{j} + \frac{\partial F}{\partial z}\vec{k}. \tag{1}$$

Il se pose la question de savoir sous quelles conditions un champ vectoriel $\vec{a}(M)$ est potentiel. Cette question a été en fait envisagée au § 7. Soit

$$\vec{a} = \vec{a}(M) = P\vec{i} + Q\vec{j} + R\vec{k}.$$

D'après la relation (1), le champ vectoriel $\vec{a}(M)$ est potentiel s'il existe une fonction $F(M)$ telle que

$$\frac{\partial F}{\partial x} = P\,; \quad \frac{\partial F}{\partial y} = Q\,; \quad \frac{\partial F}{\partial z} = R. \tag{2}$$

Dans le théorème 13.7 on a montré que l'expression $P\,dx + Q\,dy + R\,dz$ (où P, Q et R sont des fonctions pourvues de dérivées partielles premières continues) est la différentielle totale d'une fonction $F(x, y, z)$ si et seulement si P, Q et R satisfont aux conditions

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}\,; \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}\,; \quad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z}\,. \tag{3}$$

Or si $P\,dx + Q\,dy + R\,dz = dF$, on a les relations (2), c'est-à-dire que la condition (3) exprime précisément que le champ vectoriel envisagé est potentiel. La fonction $F(x, y, z)$ s'appelle alors *fonction potentielle* du champ.

Un exemple de champ potentiel nous est fourni par le champ de gravitation. Une masse m placée en l'origine des coordonnées

engendre un champ de gravitation; chaque point M de l'espace affecté d'une masse unité est soumis d'après la loi de Newton à l'action d'une force $\vec{F}(M)$ d'intensité $k\frac{m}{r^2}$ dirigée vers l'origine des coordonnées. Ici $r = |\overrightarrow{OM}| = \sqrt{x^2+y^2+z^2}$, k est le coefficient de proportionnalité.

Soit $M = M(x; y; z)$. Les composantes de la force $\vec{F}(M)$ se définissent comme suit :

$$P = |\vec{F}| \cos\alpha = \frac{km}{r^2}\left(-\frac{x}{r}\right) = -\frac{kmx}{r^3},$$

$$Q = |\vec{F}| \cos\beta = \frac{km}{r^2}\left(-\frac{y}{r}\right) = -\frac{kmy}{r^3},$$

$$R = |\vec{F}| \cos\gamma = \frac{km}{r^2}\left(-\frac{z}{r}\right) = -\frac{kmz}{r^3},$$

où $\cos\alpha$, $\cos\beta$ et $\cos\gamma$ sont les cosinus directeurs du vecteur $\vec{F}(M)$. Donc

$$\vec{F}(M) = -\frac{kmx}{r^3}\vec{i} - \frac{kmy}{r^3}\vec{j} - \frac{kmz}{r^3}\vec{k}.$$

On vérifie que ce champ vectoriel est potentiel et sa fonction potentielle est égale à $u(r) = \frac{km}{r}$.

Calculons en conclusion le travail fourni par la force $\vec{F}(M)$ pour déplacer une masse unité de $B(x_1; y_1; z_1)$ en $C(x_2; y_2; z_2)$.

On sait que le travail A s'exprime au moyen de l'intégrale curviligne

$$A = \int_{\overset{\frown}{BC}} P\,dx + Q\,dy + R\,dz,$$

où P, Q et R sont les composantes de la force $\vec{F}(M)$. Ce champ de forces étant potentiel, l'intégrant est une différentielle totale, donc l'intégrale ne dépend pas du chemin d'intégration et peut être calculée à l'aide de la formule

$$A = \int_B^C P\,dx + Q\,dy + R\,dz = u(C) - u(B),$$

c'est-à-dire que le travail de la force $\vec{F}(M)$ est égal à la différence des valeurs de la fonction potentielle aux points C et B. Ici

$$A = u(r_2) - u(r_1) = \frac{km}{r_2} - \frac{km}{r_1} = km\left(\frac{1}{r_2} - \frac{1}{r_1}\right),$$

où $r_1 = OB$ et $r_2 = OC$.

Remarquons que le domaine dans lequel est défini le champ de gravitation est l'espace tout entier privé de l'origine des coordonnées.

4. Flux d'un champ vectoriel. Soit donné dans l'espace un champ de vitesse $\vec{v}(M)$ d'un liquide, autrement dit, l'espace est rempli d'un liquide en mouvement dont la vitesse en chaque point M est définie par le vecteur

$$\vec{v}(M) = P(x, y, z)\vec{i} + Q(x, y, z)\vec{j} + R(x, y, z)\vec{k}.$$

Supposons que les fonctions P, Q et R sont continues. Calculons la quantité Π de liquide traversant pendant l'unité de temps une surface orientée S limitée par une courbe gauche L, en admettant que la densité du liquide est $\rho = 1$.

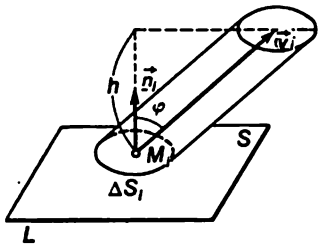


Fig. 213

Soit

$$\vec{n} = \cos\alpha\vec{i} + \cos\beta\vec{j} + \cos\gamma\vec{k}$$

le vecteur normal unitaire à la surface S et supposons que ses cosinus directeurs sont des fonctions continues des coordonnées x, y, z des points de S.

Partageons la surface S arbitrairement en n parties d'aires ΔS_1, ΔS_2, . . ., ΔS_n et choisissons un point $M_i(x_i; y_i; z_i)$ dans chacune d'elles. Trouvons la quantité Π_i de liquide traversant pendant l'unité de temps la surface partielle S_i (fig. 213). Désignons par φ l'angle des vecteurs $\vec{n}_i = \vec{n}(M_i)$ et $\vec{v}_i = \vec{v}(M_i)$. Si cet angle est aigu, c'est-à-dire si le liquide traverse la surface S_i dans le sens de la normale $\vec{n}_i$, la quantité Π_i sera supposée positive, si l'angle est obtus, c'est-à-dire si le liquide coule dans le sens opposé, Π_i sera supposée négative.

On peut considérer approximativement que pour une subdivision assez fine de la surface S, la vitesse $\vec{v}$ est constante en tous les points de S_i et est égale à $\vec{v}(M_i)$, et que les surfaces partielles sont planes. La quantité Π_i est alors approximativement égale au volume, convenablement signé, du cylindre de base ΔS_i et de hauteur égale au module du projeté du vecteur $\vec{v}_i$ sur la normale $\vec{n}_i$, c'est-à-dire que $\Pi_i \approx \Delta S_i h$, où h est la hauteur indiquée.

Comme $h = |\vec{v}_i| \cos\varphi = |\vec{v}_i|\,|\vec{n}_i| \cos\varphi = (\vec{v}_i \cdot \vec{n}_i)$, il vient

$$\Pi_i \approx (\vec{v}_i \cdot \vec{n}_i)\,\Delta S_i.$$

En sommant sur i entre 1 et n, on trouve une valeur approchée de la quantité Π de liquide traversant la surface orientée S pendant l'unité de temps:

$$\Pi \approx \sum_{i=1}^{n} (\vec{v}_i \cdot \vec{n}_i)\,\Delta S_i.$$

La somme de droite est la somme intégrale de la fonction $(\vec{v} \cdot \vec{n})$. Les composantes P, Q et R du vecteur $\vec{v}$ et les cosinus directeurs du vecteur $\vec{n}$ étant des fonctions continues de x, y, z, il en est de même du produit scalaire $\vec{v} \cdot \vec{n} = P \cos\alpha + Q \cos\beta + R \cos\gamma$. Donc lorsque le plus grand des diamètres des surfaces partielles S_i tend vers 0, cette somme admet une limite qui est égale à l'intégrale de surface de première espèce de la fonction $(\vec{v} \cdot \vec{n})$ étendue à S. Un passage à la limite nous donne la valeur exacte de Π:

$$\Pi = \lim_{\lambda \to 0} \sum_{i=1}^{n} (\vec{v}_i \cdot \vec{n}_i)\,\Delta S_i = \iint_S (\vec{v} \cdot \vec{n})\,dS,$$

ou, si l'on exprime le produit scalaire en fonction des composantes des vecteurs,

$$\Pi = \iint_S (P \cos\alpha + Q \cos\beta + R \cos\gamma)\,dS.$$

En se servant de la formule (14) du § 11 qui lie les intégrales de surface de première et de seconde espèce, on obtient finalement

$$\Pi = \iint_S P\,dy\,dz + Q\,dz\,dx + R\,dx\,dy. \qquad (4)$$

Donc la quantité Π de liquide traversant pendant l'unité de temps une surface orientée S est une intégrale de surface de seconde espèce étendue au côté correspondant de S.

Pour un champ vectoriel arbitraire $\vec{v}(M)$, l'intégrale de surface de seconde espèce (4) s'appelle *flux du vecteur* $\vec{v}(M)$ à travers la surface S. Il est entendu que si le champ vectoriel est d'une nature autre que dans l'exemple considéré, le flux admet une autre signification physique.

5. Divergence. Soit donné sur un domaine V limité par une surface bornée S un champ vectoriel $\vec{a}(M) = \{P; Q; R\}$ tel que

les fonctions $P(M)$, $Q(M)$ et $R(M)$ sont continues avec leurs dérivées partielles sur V.

Définition 1. On appelle *divergence* du champ vectoriel $\vec{a}(M)$ la fonction scalaire $\operatorname{div}\vec{a}(M)$ définie par l'égalité

$$\operatorname{div}\vec{a}(M)=\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}. \tag{5}$$

En se servant de l'expression de la divergence et de la notion de flux d'un vecteur à travers une surface, on peut écrire la formule d'Ostrogradski (formule (1), § 12) sous une forme vectorielle plus condensée. L'intégrale de surface de la formule d'Ostrogradski est le flux du vecteur $\vec{a}=\{P; Q; R\}$ à travers la surface S:

$$\iint_S P\,dy\,dz+Q\,dz\,dx+R\,dx\,dy=\iint_S (\vec{a}\cdot\vec{n})\,dS.$$

En se servant de cette expression et de la formule (5), on peut mettre la formule d'Ostrogradski sous la forme

$$\iint_S (\vec{a}\cdot\vec{n})\,dS=\iiint_V \operatorname{div}\vec{a}(M)\,dv. \tag{6}$$

Donc le flux du vecteur $\vec{a}(M)$ à travers la surface fermée S est égal à l'intégrale triple de la divergence du champ $\vec{a}(M)$ étendue au domaine limité par S.

Montrons que la divergence est indépendante du choix du système de coordonnées, bien que sa définition fasse intervenir ce système. A cet effet, considérons un point arbitraire M, comprenons-le dans un volume V limité par une surface S et appliquons la formule d'Ostrogradski à V. En utilisant ensuite le théorème de la moyenne pour une intégrale triple, on trouve

$$\iint_S (\vec{a}\cdot\vec{n})\,dS=\operatorname{div}\vec{a}(M_0)\,v,$$

où M_0 est un point quelconque de V; v, le volume de V. D'où

$$\operatorname{div}\vec{a}(M_0)=\frac{\iint_S (\vec{a}\cdot\vec{n})\,dS}{v}.$$

Si maintenant on fait dégénérer V en M, alors $v\to 0$, $M_0\to M$ et l'on obtient

$$\operatorname{div}\vec{a}(M)=\lim_{v\to 0}\frac{\iint_S (\vec{a}\cdot\vec{n})\,dS}{v}, \tag{7}$$

autrement dit la divergence du champ vectoriel $\vec{a}(M)$ en M est égale à la limite du rapport du flux du vecteur $\vec{a}(M)$ à travers la surface S limitant le point M, au volume limité par S. Comme le flux et le volume ne dépendent pas du choix du système de coordonnées, il en est de même de la divergence. C.Q.F.D.

Déterminons maintenant la signification physique de la divergence à l'aide de la formule (7). Traitons à cet effet le champ vectoriel $\vec{a}(M)$ comme le champ de vitesse d'un liquide de densité $\rho = 1$. Comme établi au n° 4, le flux

$$\Pi = \int\int_S (\vec{a} \cdot \vec{n})\, dS$$

du vecteur $\vec{a}(M)$ est égal à la quantité de liquide traversant pendant l'unité de temps la surface S dans le sens de la normale $\vec{n}$. Supposons que $\vec{n}$ est la normale extérieure. La surface S étant fermée, il est évident que le flux du vecteur $\vec{a}(M)$ est égal à la quantité de liquide qui entre ou sort pendant l'unité de temps du volume V limité par S. Appelons cette quantité *débit total des sources* (si $\Pi > 0$) ou *des puits* (si $\Pi < 0$) situés dans V. Considérons le rapport

$$\frac{\int\int_S (\vec{a} \cdot \vec{n})\, dS}{v}.$$

Il représente la densité moyenne des sources (ou des puits), c'est-à-dire la quantité de liquide qui apparaît (ou disparaît) pendant l'unité de temps dans une unité de volume V, et la limite

$$\lim_{r \to 0} \frac{\int\int_S (\vec{a} \cdot \vec{n})\, dS}{v},$$

à condition que V dégénère en M, peut être appelée *densité* des sources (ou des puits) en M. Or, cette limite est égale à $\operatorname{div} \vec{a}(M)$. Donc la divergence d'un champ de vitesses caractérise la densité des sources de liquide.

Si $\operatorname{div} \vec{a}(M) > 0$, il résulte de la formule (6) que $\Pi > 0$, c'est-à-dire que le volume V contient des sources et il en sort plus de liquide qu'il n'y entre; si $\operatorname{div} \vec{a}(M) < 0$, alors $\Pi < 0$, c'est-à-dire que le volume V contient des puits et il entre plus de liquide qu'il n'en sort. Si $\operatorname{div} \vec{a}(M) = 0$, alors $\Pi = 0$, c'est-à-dire que le volume V ne contient ni sources ni puits, et il sort de

V autant de liquide qu'il n'y entre. C'est le cas, par exemple, de tout volume V situé dans un courant d'eau.

La divergence d'un champ vectoriel arbitraire admet la même interprétation physique : elle caractérise la densité des sources de ce champ.

On dit qu'un champ vectoriel $\vec{a}(M)$ est *solénoïdal* si $\operatorname{div} \vec{a}(M) = 0$. Comme exemple de champ solénoïdal citons le champ de vitesse d'un liquide privé de sources et de puits envisagé ci-dessus.

Exemple 2. Calculer la divergence du champ de vitesse $\vec{v}(M) = -\omega y\vec{i} + \omega x\vec{j}$ d'un solide tournant à une vitesse angulaire constante ω autour de l'axe Oz.

Solution. On a ici $P = -\omega y$, $Q = \omega x$, $R = 0$. Donc

$$\operatorname{div} \vec{v}(M) = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = \frac{\partial(-\omega y)}{\partial x} + \frac{\partial(\omega x)}{\partial y} = 0,$$

c'est-à-dire que ce champ est solénoïdal.

6. Circulation. Rotationnel. Soit défini sur un domaine un champ vectoriel

$$\vec{a}(M) = P(M)\vec{i} + Q(M)\vec{j} + R(M)\vec{k}$$

et soit L une courbe différentiable ou différentiable par morceaux située dans ce domaine. Choisissons une orientation sur L et désignons par $\vec{dl}$ le vecteur

$$\vec{dl} = dx\vec{i} + dy\vec{j} + dz\vec{k}.$$

Alors l'intégrale curviligne du produit scalaire des vecteurs $\vec{a}(M)$ et $\vec{dl}$

$$\int_L \vec{a}(M) \cdot \vec{dl} = \int_L P\,dx + Q\,dy + R\,dz$$

s'appelle *circulation du champ vectoriel* $\vec{a}(M)$ *le long de la courbe* L. Dans un champ de force la circulation exprime le travail de ce champ pour déplacer un point matériel le long du chemin L. La circulation admet une autre interprétation physique pour les champs d'une autre nature.

Définition 2. On appelle *rotationnel d'un champ vectoriel* $\vec{a}(M)$ le vecteur $\overrightarrow{\operatorname{rot}}\,\vec{a}(M)$ défini par l'égalité

$$\overrightarrow{\operatorname{rot}}\,\vec{a}(M) = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\vec{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\vec{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\vec{k}.$$

Les notions de rotationnel et de circulation nous permettent d'écrire la formule de Stokes

$$\oint_L P\,dx+Q\,dy+R\,dz=\iint_S\left[\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\cos\alpha+\right.$$
$$\left.+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\cos\beta+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\cos\gamma\right]dS$$

sous la forme vectorielle condensée:

$$\oint_L \vec{a}(M)\cdot d\vec{l}=\iint_S(\overrightarrow{\text{rot}}\,\vec{a}(M)\cdot\vec{n})\,dS.$$

Donc la circulation d'un champ vectoriel $\vec{a}(M)$ le long d'un contour fermé L est égale au flux du rotationnel de ce champ à travers la surface S limitée par le contour L.

De même que pour la divergence, on démontre que $\overrightarrow{\text{rot}}\,\vec{a}(M)$ dépend du seul champ $\vec{a}(M)$ et pas du système de coordonnées.

Exemple 3. Calculer le rotationnel du champ de vitesse $\vec{v}(M)=-\omega y\vec{i}+\omega x\vec{j}$ d'un solide tournant à une vitesse angulaire constante ω autour de l'axe Oz.

Solution. La définition du rotationnel nous donne

$$\overrightarrow{\text{rot}}\,\vec{a}(M)=\left(\frac{\partial 0}{\partial y}-\frac{\partial(\omega x)}{\partial z}\right)\vec{i}+\left(\frac{\partial(-\omega y)}{\partial z}-\frac{\partial 0}{\partial x}\right)\vec{j}+$$
$$+\left(\frac{\partial(\omega x)}{\partial x}-\frac{\partial(-\omega y)}{\partial y}\right)\vec{k}=2\omega\vec{k},$$

autrement dit, le rotationnel de ce champ est de même direction et de même sens que l'axe Oz et son module est égal au double de la vitesse angulaire.

La notion de rotationnel est directement attachée à la notion de champ potentiel. On a déjà montré qu'un champ vectoriel $\vec{a}(M)=\{P;\ Q;\ R\}$ est potentiel si et seulement si

$$\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x};\quad \frac{\partial Q}{\partial z}=\frac{\partial R}{\partial y};\quad \frac{\partial R}{\partial x}=\frac{\partial P}{\partial z}.$$

Or, ceci exprime que les composantes du rotationnel du champ $\vec{a}(M)$ sont nulles. Donc une condition nécessaire et suffisante pour qu'un champ $\vec{a}(M)$ soit potentiel est que

$$\overrightarrow{\text{rot}}\,\vec{a}(M)\equiv 0.$$

7. Opérateur de Hamilton. L'opérateur de Hamilton *) ou opérateur nabla

$$\nabla = \frac{\partial}{\partial x}\vec{i} + \frac{\partial}{\partial y}\vec{j} + \frac{\partial}{\partial z}\vec{k}.$$

se prête bien à l'étude des notions fondamentales de la théorie du champ: gradient, divergence, rotationnel, et aux opérations sur eux. On traitera l'opérateur ∇ comme un vecteur symbolique de composantes $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ et $\frac{\partial}{\partial z}$ soumis aux règles opératoires de l'algèbre vectorielle. De plus, par produit de $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ et $\frac{\partial}{\partial z}$ par une fonction scalaire, on comprendra la dérivée partielle de cette fonction respectivement par rapport à x, y et z.

Exemples.

1. Soit $u(x, y, z)$ une fonction scalaire. Le produit de l'opérateur ∇ par la fonction u nous donne le gradient de cette fonction:

$$\nabla u = \left(\frac{\partial}{\partial x}\vec{i} + \frac{\partial}{\partial y}\vec{j} + \frac{\partial}{\partial z}\vec{k}\right) u = \frac{\partial u}{\partial x}\vec{i} + \frac{\partial u}{\partial y}\vec{j} + \frac{\partial u}{\partial z}\vec{k} = \overrightarrow{\text{grad}}\, u.$$

2. Soit $\vec{a}(M) = P\vec{i} + Q\vec{j} + R\vec{z}$ une fonction vectorielle. Le produit scalaire de l'opérateur ∇ par la fonction vectorielle $\vec{a}(M)$ est la divergence de cette fonction:

$$\nabla \cdot \vec{a}(M) = \left(\frac{\partial}{\partial x}\vec{i} + \frac{\partial}{\partial y}\vec{j} + \frac{\partial}{\partial z}\vec{k}\right) \cdot (P\vec{i} + Q\vec{j} + R\vec{k}) =$$

$$= \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = \text{div}\, \vec{a}(M).$$

3. Le produit vectoriel de l'opérateur ∇ et de la fonction vectorielle $\vec{a}(M)$ est égal au rotationnel de cette fonction:

$$\nabla \wedge \vec{a}(M) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} =$$

$$= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\vec{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\vec{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\vec{k} = \overrightarrow{\text{rot}}\, \vec{a}(M).$$

Dans les applications on rencontre souvent les combinaisons deux à deux de div, $\overrightarrow{\text{rot}}$ et $\overrightarrow{\text{grad}}$. Voyons les plus importantes d'entre elles.

*) Hamilton William Rowan (1805-1865), mathématicien et astronome irlandais.

1° $\operatorname{div}\overrightarrow{\operatorname{rot}}\,\vec{a}\,(M)=0$. En effet,

$$\operatorname{div}\overrightarrow{\operatorname{rot}}\,\vec{a}\,(M)=\frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)+\frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)+\frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)=$$

$$=\frac{\partial^2 R}{\partial x\,\partial y}-\frac{\partial^2 Q}{\partial x\,\partial z}+\frac{\partial^2 P}{\partial y\,\partial z}-\frac{\partial^2 R}{\partial y\,\partial x}+\frac{\partial^2 Q}{\partial z\,\partial x}-\frac{\partial^2 P}{\partial z\,\partial y}=0,$$

puisque les dérivées mixtes secondes sont égales. On retrouve facilement ce résultat avec l'opérateur ∇:

$$\operatorname{div}\overrightarrow{\operatorname{rot}}\,\vec{a}\,(M)=\nabla\cdot(\nabla\wedge\vec{a})=0,$$

puisque nous avons affaire au produit mixte de trois « vecteurs »: ∇, ∇ et $\vec{a}$, dont deux sont identiques. Or, un tel produit est visiblement nul.

2° $\overrightarrow{\operatorname{rot}}\,\overrightarrow{\operatorname{grad}}\,u=0$. En effet,

$$\overrightarrow{\operatorname{rot}}\,\overrightarrow{\operatorname{grad}}\,u=\left[\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial z}\right)-\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial y}\right)\right]\vec{i}+\left[\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial x}\right)-\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial z}\right)\right]\vec{j}+$$

$$+\left[\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right)-\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}\right)\right]\vec{k}=\left(\frac{\partial^2 u}{\partial y\,\partial z}-\frac{\partial^2 u}{\partial z\,\partial y}\right)\vec{i}+\left(\frac{\partial^2 u}{\partial z\,\partial x}-\frac{\partial^2 u}{\partial x\,\partial z}\right)\vec{j}+$$

$$+\left(\frac{\partial^2 u}{\partial x\,\partial y}-\frac{\partial^2 u}{\partial y\,\partial x}\right)\vec{k}=0,$$

puisque les dérivées mixtes secondes sont égales. On retrouve facilement ce résultat à l'aide de l'opérateur ∇:

$$\overrightarrow{\operatorname{rot}}\,\overrightarrow{\operatorname{grad}}\,u=\nabla\wedge(\nabla u)=(\nabla\wedge\nabla)\,u=0,$$

puisque le produit vectoriel d'un vecteur par lui-même est nul.

3° $\operatorname{div}\overrightarrow{\operatorname{grad}}\,u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}$. En effet,

$$\overrightarrow{\operatorname{grad}}\,u=\frac{\partial u}{\partial x}\vec{i}+\frac{\partial u}{\partial y}\vec{j}+\frac{\partial u}{\partial z}\vec{k},$$

$$\operatorname{div}\overrightarrow{\operatorname{grad}}\,u=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\right)+$$

$$+\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial z}\right)=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}. \tag{8}$$

Le second membre de (8) se désigne symboliquement par

$$\Delta u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}\quad\text{ou}\quad\Delta u=\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)u.$$

Le symbole

$$\Delta=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$$

s'appelle *opérateur de Laplace* *). Il est naturel de traiter l'opérateur de Laplace Δ comme le carré scalaire de ∇. En effet,

$$(\nabla \cdot \nabla) = \left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 + \left(\frac{\partial}{\partial z}\right)^2 = \Delta.$$

Donc l'opérateur ∇ permet d'écrire l'égalité (8) sous la forme

$$\operatorname{div} \overrightarrow{\operatorname{grad}}\, u = \nabla \cdot (\nabla u) = (\nabla \cdot \nabla)\, u = \nabla^2 u.$$

L'équation

$$\Delta u = 0$$

s'appelle *équation de Laplace*. On l'utilise pour décrire divers processus stationnaires tels la distribution stationnaire de la chaleur, le champ électrostatique des charges ponctuelles, le mouvement permanent de fluide incompressible à l'intérieur d'un domaine, etc. Un champ scalaire $u(x, y, z)$ vérifiant la condition $\Delta u = 0$ s'appelle *laplacien* ou *harmonique*.

---

*) Laplace Pierre Simon (1749-1827), mathématicien et physicien français.

TROISIÈME PARTIE

SÉRIES, ÉQUATIONS DIFFÉRENTIELLES

CHAPITRE XIV

SÉRIES

Dans ce chapitre nous étudions les séries qui constituent un important outil de calcul et de recherche aussi bien en mathématiques que dans nombre de ses applications.

§ 1. Notion de série numérique

1. Définitions fondamentales. Soit donnée une suite numérique $a_1, a_2, \ldots, a_n, \ldots$ L'expression

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots = \sum_{n=1}^{\infty} a_n \qquad (1)$$

s'appelle *série numérique* ou simplement *série.*

Les nombres $a_1, a_2, \ldots, a_n, \ldots$ sont les *termes* ou les *éléments* de la série, a_n est le *terme général* de la série.

Les sommes d'un nombre fini de termes de la série :

$$S_1 = a_1, \quad S_2 = a_1 + a_2, \quad S_3 = a_1 + a_2 + a_3, \ldots,$$
$$S_n = a_1 + a_2 + a_3 + \ldots + a_n, \ldots$$

s'appellent *sommes partielles* de la série (1). Les sommes partielles forment une suite infinie

$$S_1, S_2, \ldots, S_n, \ldots \qquad (2)$$

On dit que la série (1) *converge* si la suite de ses sommes partielles (2) converge vers un nombre S qui s'appelle alors *somme de la série* (1). Ce qu'on note symboliquement

$$S = a_1 + a_2 + \ldots + a_n + \ldots \quad \text{ou} \quad S = \sum_{n=1}^{\infty} a_n.$$

Si la suite des sommes partielles (2) diverge, on dit que la série (1) *diverge.*

Exemple 1. Montrons que la série

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \ldots + \frac{1}{n(n+1)} + \ldots = \sum_{n=1}^{\infty} \frac{1}{n(n+1)}$$

converge. Considérons la somme partielle S_n

$$S_n = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \ldots + \frac{1}{n(n+1)}.$$

Les termes de cette somme peuvent être mis sous la forme

$$\frac{1}{1\cdot 2} = 1 - \frac{1}{2}; \quad \frac{1}{2\cdot 3} = \frac{1}{2} - \frac{1}{3};$$

$$\frac{1}{3\cdot 4} = \frac{1}{3} - \frac{1}{4}; \ \ldots; \quad \frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}.$$

Donc

$$S_n = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) +$$

$$+ \left(\frac{1}{3} - \frac{1}{4}\right) + \ldots + \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{n+1}.$$

D'où il s'ensuit que la limite de la suite des sommes partielles est égale à 1 :

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \left(1 - \frac{1}{n+1}\right) = 1 - \lim_{n\to\infty} \frac{1}{n+1} = 1.$$

Donc la série converge et sa somme S est égale à 1.

Exemple 2. Etudions la série

$$1 - 1 + 1 - 1 + \ldots + (-1)^{n-1} + \ldots = \sum_{n=1}^{\infty} (-1)^{n-1}.$$

La suite de ses sommes partielles est $S_1 = 1$, $S_2 = 0$, $S_3 = 1$, $S_4 = 0$, . . . Cette suite ne tend vers aucune limite, donc la série envisagée diverge.

Exemple 3. Etudions la série

$$a + aq + aq^2 + \ldots + aq^{n-1} + \ldots = \sum_{n=1}^{\infty} aq^{n-1}, \ a \neq 0. \qquad (3)$$

On reconnaît une progression géométrique de raison q dont la n-ième somme partielle S_n est

$$S_n = a + aq + aq^2 + \ldots + aq^{n-1} = \frac{a - aq^n}{1-q} = \frac{a}{1-q} - \frac{aq^n}{1-q}.$$

Donc

1) si $|q| < 1$, alors $\lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty} \frac{a}{1-q} - \lim\limits_{n\to\infty} \frac{aq^n}{1-q} = \frac{a}{1-q}$,

c'est-à-dire que la série converge et sa somme $S = \frac{a}{1-q}$. Pour $a = 1$ et $q = 1/2$ par exemple, on a

$$S = 1 + \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}} + \ldots = 2;$$

2) si $|q|>1$, alors $\lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty} \frac{a-aq^n}{1-q} = \infty$, et la série diverge;

3) si $q=1$, la série (3) devient $a+a+\ldots+a+\ldots$ Dans ce cas $\lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty} n\cdot a = \infty$ et la série diverge;

4) si $q=-1$, la série (3) devient $a-a+a-a+\ldots$ On a alors $S_n = \frac{a}{2} - \frac{a(-1)^n}{2}$, c'est-à-dire que $S_n=0$ pour n pair et $S_n=a$ pour n impair. Donc $\lim\limits_{n\to\infty} S_n$ n'existe pas et la série diverge.

En résumé, la série (3) converge pour $|q|<1$ et diverge pour $|q|\geqslant 1$.

2. Propriétés des séries convergentes.

Théorème 14.1. *Si la série*

$$a_1+a_2+a_3+\ldots+a_{k-1}+a_k+a_{k+1}+\ldots$$

$$\ldots+a_{n-1}+a_n+\ldots=\sum_{n=1}^{\infty} a_n, \qquad (4)$$

converge, il en est de même de la série

$$a_{k+1}+\ldots+a_{n-1}+a_n+\ldots=\sum_{n=k+1}^{\infty} a_n, \qquad (5)$$

et réciproquement.

En d'autres termes, la convergence d'une série n'est pas affectée par l'élimination d'un nombre fini de premiers termes.

Démonstration. Supposons que la série (4) converge et soit S sa somme: $\lim\limits_{n\to\infty} S_n = S$. Désignons par S_k la somme des termes éliminés dans (4) et par σ_{n-k} la somme des $n-k$ premiers termes de la série (5). Alors

$$S_n = S_k + \sigma_{n-k}, \qquad (6)$$

où S_k est un nombre indépendant de n. L'égalité (6) entraîne

$$\lim_{n\to\infty} \sigma_{n-k} = \lim_{n\to\infty} (S_n - S_k) = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_k = S - S_k,$$

ce qui exprime que la suite des sommes partielles $\{\sigma_{n-k}\}$ de la série (5) admet une limite, donc que cette série converge.

Supposons maintenant que la série (5) converge et soit σ sa somme: $\lim\limits_{n\to\infty} \sigma_{n-k} = \sigma$. De (6) il s'ensuit alors que

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} (S_k + \sigma_{n-k}) = \lim_{n\to\infty} S_k + \lim_{n\to\infty} \sigma_{n-k} = S_k + \sigma,$$

ce qui exprime la convergence de la série (4). ■

Les séries convergentes sont justiciables des opérations arithmétiques usuelles.

Théorème 14.2. *Si la série* $\sum_{n=1}^{\infty} a_n$ *converge et sa somme est égale à* S, *la série* $\sum_{n=1}^{\infty} ca_n$ *où* c *est une constante, converge aussi et sa somme est égale à* cS.

Démonstration. Soient S_n la n-ième somme partielle de la série $\sum_{n=1}^{\infty} a_n$ et σ_n celle de la série $\sum_{n=1}^{\infty} ca_n$. Alors

$$\sigma_n = ca_1 + ca_2 + ca_3 + \ldots + ca_n = c(a_1 + a_2 + a_3 + \ldots + a_n) = cS_n.$$

En faisant tendre n vers ∞, on obtient

$$\lim_{n\to\infty} \sigma_n = \lim_{n\to\infty} cS_n = c \lim_{n\to\infty} S_n = cS,$$

c'est-à-dire que la suite des sommes partielles σ_n de la série $\sum_{n=1}^{\infty} ca_n$ converge vers cS. Donc $\sum_{n=1}^{\infty} ca_n = cS$. ■

Théorème 14.3. *Si les séries* $\sum_{n=1}^{\infty} a_n$ *et* $\sum_{n=1}^{\infty} b_n$ *convergent et leurs sommes sont respectivement égales à* S *et* σ, *la série* $\sum_{n=1}^{\infty} (a_n \pm b_n)$ *converge aussi et sa somme est égale à* $S \pm \sigma$.

Démonstration. Soient S_n et σ_n les n-ièmes sommes partielles des séries $\sum_{n=1}^{\infty} a_n$ et $\sum_{n=1}^{\infty} b_n$, et τ_n celle de la série $\sum_{n=1}^{\infty} (a_n \pm b_n)$. Alors

$$\tau_n = (a_1 \pm b_1) + (a_2 \pm b_2) + \ldots + (a_n \pm b_n) =$$
$$= (a_1 + a_2 + \ldots + a_n) \pm (b_1 + b_2 + \ldots + b_n) = S_n \pm \sigma_n.$$

En faisant tendre n vers ∞, on trouve

$$\lim_{n\to\infty} \tau_n = \lim_{n\to\infty} (S_n \pm \sigma_n) = \lim_{n\to\infty} S_n \pm \lim_{n\to\infty} \sigma_n = S \pm \sigma,$$

c'est-à-dire que la suite des sommes partielles $\{\tau_n\}$ de la série $\sum_{n=1}^{\infty} (a_n \pm b_n)$ tend vers $S + \sigma$. Donc $\sum_{n=1}^{\infty} (a_n \pm b_n) = S \pm \sigma$. ■

On a ainsi montré que les séries convergentes peuvent être multipliées par un nombre, additionnées et soustraites terme à terme comme les sommes finies.

3. Condition nécessaire de convergence d'une série. Etudier une série, c'est : 1) dire si elle converge ; 2) sachant qu'elle converge, trouver sa somme. On se penchera essentiellement sur le premier problème qui est de nature théorique. Exhibons une condition nécessaire de convergence d'une série.

Théorème 14.4. *Si la série* $\sum_{n=1}^{\infty} a_n$ *converge, son terme général tend vers* 0, *c'est-à-dire que* $\lim_{n\to\infty} a_n = 0$.

Démonstration. La série $\sum_{n=1}^{\infty} a_n$ converge par hypothèse. Désignons sa somme par S. Considérons les sommes partielles $S_n = a_1 + a_2 + \ldots + a_{n-1} + a_n$ et $S_{n-1} = a_1 + a_2 + \ldots + a_{n-1}$. D'où $a_n = S_n - S_{n-1}$. Puisque $S_n \to S$ et $S_{n-1} \to S$ lorsque $n \to \infty$, on a

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (S_n - S_{n-1}) = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = S - S = 0. \blacksquare$$

La condition $\lim_{n\to\infty} a_n = 0$ est une condition nécessaire mais pas suffisante de convergence de la série.

Exemple. Considérons la série dite harmonique

$$1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots = \sum_{n=1}^{\infty} \frac{1}{n}.$$

Il est évident que la série harmonique est justiciable de la condition nécessaire de convergence, puisque $\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{1}{n} = 0$. Montrons que cette série diverge. En effet, si elle convergeait, en désignant par S sa somme, on aurait

$$\lim_{n\to\infty} (S_{2n} - S_n) = \lim_{n\to\infty} S_{2n} - \lim_{n\to\infty} S_n = S - S = 0.$$

Or,

$$S_{2n} - S_n = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} > \frac{1}{2n} + \frac{1}{2n} + \ldots + \frac{1}{2n} =$$

$$= n\,\frac{1}{2n} = \frac{1}{2},$$

donc $S_{2n} - S_n > 1/2$. D'où il suit que $\lim_{n\to\infty} (S_{2n} - S_n) = 0$, ce qui est impossible. Cette contradiction prouve que la série diverge.

Donc si le terme général d'une série tend vers 0, on ne peut se prononcer sur la convergence de cette série. Une étude complémentaire s'impose, étude qui peut être effectuée à l'aide des conditions suffisantes de convergence.

Si le terme général d'une série ne tend pas vers 0, le théorème 14.4 nous permet d'affirmer immédiatement que cette série diverge.

§ 2. Séries à termes positifs

Passons maintenant à l'étude de quelques conditions suffisantes de convergence des séries à termes positifs. Prouvons préalablement un théorème qui nous sera utile dans la suite.

Théorème 14.5. *Pour qu'une série à termes positifs* $\sum_{n=1}^{\infty} a_n$ *converge, il est nécessaire et suffisant que la suite de ses sommes partielles soit bornée.*

Démonstration. *Nécessité.* Supposons que la série $\sum_{n=1}^{\infty} a_n$ converge. Ceci exprime que la suite de ses sommes partielles possède une limite. Or, d'après le théorème 2.6, toute suite convergente est bornée.

Suffisance. Supposons que la suite des sommes partielles de la série $\sum_{n=1}^{\infty} a_n$ est bornée. Cette série étant à termes positifs, ses sommes partielles forment une suite croissante : $0 \leqslant S_1 \leqslant S_2 \leqslant \ldots \leqslant S_n \leqslant S_{n+1} \leqslant \ldots$ Cette suite converge d'après le théorème 2.12 des suites monotones bornées. Donc la série $\sum_{n=1}^{\infty} a_n$ converge. ■

Conditions suffisantes de convergence d'une série. Etablissons maintenant quelques critères de convergence des séries.

Critère de comparaison.

Théorème 14.6. *Soient données deux séries à termes positifs* $\sum_{n=1}^{\infty} a_n$ *et* $\sum_{n=1}^{\infty} b_n$ *telles que* $a_n \leqslant b_n$ *pour tout n. Alors la convergence de la série* $\sum_{n=1}^{\infty} b_n$ *entraîne celle de la série* $\sum_{n=1}^{\infty} a_n$ *et la divergence de la série* $\sum_{n=1}^{\infty} a_n$, *celle de la série* $\sum_{n=1}^{\infty} b_n$.

Démonstration. Soient S_n et σ_n les n-ièmes sommes partielles respectives des séries $\sum_{n=1}^{\infty} a_n$ et $\sum_{n=1}^{\infty} b_n$. Les inégalités

$a_n \leqslant b_n$ entraînent

$$S_n \leqslant \sigma_n. \tag{7}$$

Si la série $\sum_{n=1}^{\infty} b_n$ converge, la suite de ses sommes partielles est bornée en vertu de la condition nécessaire du théorème 14.5. c'est-à-dire que pour tout n on a $\sigma_n \leqslant M$, où M est un nombre. D'après la formule (7) on a alors $S_n \leqslant M$, ce qui exprime, en vertu de la condition suffisante du théorème 14.5, que la série $\sum_{n=1}^{\infty} a_n$ converge.

Si la série $\sum_{n=1}^{\infty} a_n$ diverge, il en est de même de la série $\sum_{n=1}^{\infty} b_n$. En effet, si l'on admet que la série $\sum_{n=1}^{\infty} b_n$ converge, il en sera de même de la série $\sum_{n=1}^{\infty} a_n$ d'après ce qui vient d'être démontré. Ce qui est contraire à l'hypothèse du théorème. ■

Exemple 1. La série $\sum_{n=1}^{\infty} \frac{1}{(n+1)^{n+1}}$ converge puisque ses termes sont majorés par les termes de la progression géométrique convergente $\sum_{n=1}^{\infty} \frac{1}{2^{n+1}} \left(q = \frac{1}{2} < 1\right)$: $\frac{1}{(n+1)^{n+1}} \leqslant \frac{1}{2^{n+1}}$.

Exemple 2. La série $\sum_{n=1}^{\infty} \frac{1}{n^s} (s \leqslant 1)$ diverge, puisque ses termes majorent les termes de la série harmonique $\sum_{n=1}^{\infty} \frac{1}{n}$ qui, on le sait, diverge : $\frac{1}{n} \leqslant \frac{1}{n^s}$.

Il existe des critères de convergence qui permettent de dire directement si une série converge ou non sans la comparer à une autre. Etudions deux d'entre eux.

Critère de d'Alembert *).

Théorème 14.7. *Soit donnée une série* $\sum_{n=1}^{\infty} a_n$ *à termes positifs et supposons que* $\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = \rho$. *Alors*

*) D'Alembert Jean Le Rond (1717-1783), mathématicien et philosophe français.

a) *si* $\rho < 1$, *cette série converge*;

b) *si* $\rho > 1$, *cette série diverge.*

Démonstration. a) Soit $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} = \rho < 1$. Montrons que la série $\sum\limits_{n=1}^{\infty} a_n$ converge. Par définition de la limite d'une suite numérique, pour tout $\varepsilon > 0$ il existe un rang N à partir duquel $(n \geqslant N) \left| \frac{a_{n+1}}{a_n} - \rho \right| < \varepsilon$. D'où

$$\rho - \varepsilon < \frac{a_{n+1}}{a_n} < \rho + \varepsilon. \tag{8}$$

Puisque $\rho < 1$, on peut prendre ε assez petit pour que $\rho + \varepsilon < 1$. En posant $\rho + \varepsilon = q$, on obtient en vertu de l'inégalité (8) de droite

$$\frac{a_{n+1}}{a_n} < q \quad \text{ou} \quad a_{n+1} < a_n q$$

pour $n = N,\ N+1,\ N+2,\ \ldots$ En attribuant ces valeurs à n, on déduit de la dernière inégalité que

$$\begin{aligned} a_{N+1} &< a_N q, \\ a_{N+2} &< a_{N+1} q < a_N q^2, \\ a_{N+3} &< a_{N+2} q < a_N q^3, \\ &\ldots\ldots\ldots\ldots \end{aligned}$$

c'est-à-dire que les termes de la série

$$a_{N+1} + a_{N+2} + a_{N+3} + \ldots \tag{9}$$

sont majorés par les termes correspondants de la progression géométrique

$$a_N q + a_N q^2 + a_N q^3 + \ldots \tag{10}$$

Cette série (10) converge, puisque $q < 1$ (cf. exemple 3, § 1). La série (9) converge aussi d'après le critère de comparaison. Or, la série (9) a été obtenue à partir de la série $\sum\limits_{n=1}^{\infty} a_n$ par élimination d'un nombre fini de premiers termes, donc la série $\sum\limits_{n=1}^{\infty} a_n$ converge d'après le théorème 14.1.

b) Supposons maintenant que $\rho > 1$. Montrons que la série $\sum\limits_{n=1}^{\infty} a_n$ diverge. Prenons ε assez petit pour que $\rho - \varepsilon > 1$. Alors à partir d'un certain rang N $(n \geqslant N)$, en vertu de l'inégalité (8)

de gauche, on aura $\frac{a_{n+1}}{a_n} > 1$ ou $a_{n+1} > a_n$. Donc les termes de la série croissent avec leur rang à partir d'un certain N, c'est-à-dire que le terme général a_n ne tend pas vers 0 lorsque $n \to \infty$. Donc la série $\sum_{n=1}^{\infty} a_n$ diverge d'après le théorème 14.4. ■

Remarque. Pour $\rho = 1$, la série peut converger ou diverger. Dans ce cas il faut se tourner vers le critère de comparaison ou vers d'autres critères.

Exemple 3. La série $\sum_{n=1}^{\infty} \frac{1}{n!}$ converge, puisque

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{n!}{(n+1)!} = \lim_{n\to\infty} \frac{1}{n+1} = 0 < 1.$$

Exemple 4. La série $\sum_{n=1}^{\infty} \frac{n^n}{n!}$ diverge, puisque

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{(n+1)^{n+1} n!}{(n+1)!\, n^n} = \lim_{n\to\infty} \left(\frac{n+1}{n}\right)^n$$

$$= \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e > 1.$$

Exemple 5. Considérons la série $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$. On a $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} =$

$= \lim_{n\to\infty} \frac{\sqrt{n}}{\sqrt{n+1}} = 1$. Le critère de d'Alembert ne dit pas si cette série converge ou non. Mais on a établi précédemment (cf. exemple 2) que cette série diverge.

Critère intégral.

Théorème 14.8. *Soit donnée la série*

$$f(1) + f(2) + f(3) + \ldots + f(n) + \ldots = \sum_{n=1}^{\infty} f(n)$$

dont les termes sont les valeurs d'une fonction $f(x)$ continue strictement positive et strictement décroissante sur l'intervalle $[1, +\infty[$. *Alors la série* $\sum_{n=1}^{\infty} f(n)$ *est convergente ou divergente en même temps que l'intégrale* $\int_1^{+\infty} f(x)\, dx$.

Démonstration. Considérons le trapèze curviligne limité par la courbe représentative de la fonction $y = f(x)$, les droites $x = 1$, $x = n$ et l'axe Ox. Construisons les rectangles de bases $[1, 2]$, $[2, 3]$, ..., $[n-1, n]$ et de hauteurs $f(1)$, $f(2)$, $f(3)$, ..., $f(n-1)$, $f(n)$ (fig. 214). D'après la signification géométrique de l'intégrale, on a

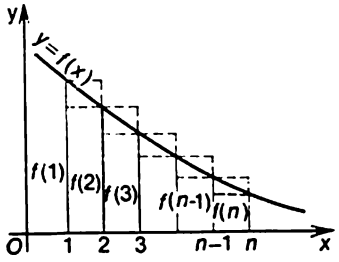


Fig. 214

$$f(2)+f(3)+\ldots+f(n)<\int_1^n f(x)\,dx<f(1)+f(2)+\ldots+f(n-1),$$

ou, de façon plus condensée,

$$S_n-f(1)<\int_1^n f(x)\,dx<S_n-f(n).$$

D'où

$$S_n<f(1)+\int_1^n f(x)\,dx, \tag{11}$$

$$S_n>f(n)+\int_1^n f(x)\,dx, \tag{12}$$

où S_n sont les sommes partielles de la série envisagée.

Supposons que l'intégrale $\int_1^{+\infty} f(x)\,dx$ converge. Ceci exprime que $\lim\limits_{n\to\infty}\int_1^n f(x)\,dx = I$. Puisque $f(x)>0$, la suite $\int_1^n f(x)\,dx$ croît avec n et est majorée par sa limite : $\int_1^n f(x)\,dx<I$. L'inégalité (11) entraîne $S_n<f(1)+I$, c'est-à-dire que la suite des sommes partielles $\{S_n\}$ de la série $\sum\limits_{n=1}^{\infty} f(n)$ est bornée supérieurement. Donc la série $\sum\limits_{n=1}^{\infty} f(n)$ converge.

Supposons maintenant que l'intégrale $\int_1^{+\infty} f(x)\,dx$ diverge. Dans ce cas $\int_1^n f(x)\,dx \to +\infty$ lorsque $n \to \infty$ (en tant que suite non bornée strictement croissante). De l'inégalité (12) il s'ensuit que $S_n \to +\infty$ lorsque $n \to \infty$, c'est-à-dire que la suite des sommes partielles $\{S_n\}$ de la série $\sum_{n=1}^{\infty} f(n)$ diverge, donc cette série aussi. ■

Exemple 6. Considérons la série $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ $(\alpha > 0)$. Etudions sa convergence à l'aide du critère intégral pour $\alpha > 0$. Prenons pour fonction $f(x)$ la fonction $\frac{1}{x^\alpha}$ $(x \geqslant 1)$ qui satisfait aux conditions du théorème 14.8. Les termes de cette série sont égaux aux valeurs de cette fonction pour $x = 1, 2, 3, \ldots, n, \ldots$ On sait (cf. chap. VIII, § 11, n°1, exemple 4) que l'intégrale impropre $\int_1^{+\infty} \frac{dx}{x^\alpha}$ converge pour $\alpha > 1$ et diverge pour $\alpha \leqslant 1$. Donc il en est de même de la série envisagée.

Remarquons que les séries de cette nature divergent aussi pour $\alpha \leqslant 0$, puisque le terme général ne tend pas vers 0 lorsque $n \to \infty$, c'est-à-dire que la condition nécessaire de convergence est violée (cf. théorème 14.4).

En particulier, la série $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ converge pour $\alpha = 2$ et diverge pour $\alpha = 1$, $\alpha = \frac{1}{2}$, etc.

§ 3. Séries alternées

Nous avons étudié les séries à termes positifs. Les séries à termes négatifs se distinguent des précédentes par leur signe seulement, donc elles sont justiciables des mêmes critères de convergence.

Penchons-nous maintenant sur les séries alternées, c'est-à-dire les séries dont les termes sont alternativement positifs ou négatifs. Pour la commodité on admettra que le premier terme est posi-

tif. Une série alternée est de la forme

$$a_1 - a_2 + a_3 - a_4 + \ldots + (-1)^{n+1} a_n + \ldots, \qquad (1)$$

où $a_n > 0$.

Pour les séries alternées on a une condition suffisante de convergence très simple.

Critère de Leibniz.

Théorème 14.9. *Si les modules des termes d'une série alternée* (1) *forment une suite strictement décroissante et le terme général de la série* (1) *tend vers* 0, *alors cette série converge.*

Démonstration. Soit donnée une série (1) et supposons que $a_n > a_{n+1}$, $a_n \to 0$ lorsque $n \to \infty$.

Considérons la somme partielle d'un nombre pair de termes

$$S_{2n} = a_1 - a_2 + a_3 - a_4 + \ldots + a_{2n-1} - a_{2n} =$$
$$= (a_1 - a_2) + (a_3 - a_4) + \ldots + (a_{2n-1} - a_{2n}).$$

Toutes les parenthèses sont strictement positives en vertu de la première condition, donc la suite des sommes partielles $\{S_{2n}\}$ est strictement croissante. Montrons qu'elle est majorée. A cet effet, mettons-la sous la forme

$$S_{2n} = a_1 - [(a_2 - a_3) + (a_4 - a_5) + \ldots +$$
$$\ldots + (a_{2n-2} - a_{2n-1}) + a_{2n}].$$

D'où il résulte que $S_{2n} < a_1$ pour tout n, c'est-à-dire que $\{S_{2n}\}$ est majorée.

La suite $\{S_{2n}\}$ étant strictement croissante et majorée, elle admet une limite S lorsque $n \to \infty$.

Montrons maintenant que la suite des sommes partielles d'un nombre impair de termes tend vers la même limite S. En effet, $S_{2n+1} = S_{2n} + a_{2n+1}$. En faisant tendre n vers ∞ et utilisant la deuxième condition ($a_n \to 0$ lorsque $n \to \infty$), on obtient

$$\lim_{n\to\infty} S_{2n+1} = \lim_{n\to\infty} (S_{2n} + a_{2n+1}) = \lim_{n\to\infty} S_{2n} + \lim_{n\to\infty} a_{2n+1} = S + 0 = S.$$

Donc la suite des sommes partielles $\{S_n\}$ de la série (1) tend vers S. Ce qui exprime que la série (1) converge. ■

Exemple. La série $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + (-1)^{n+1} \frac{1}{n} + \ldots$ converge, puisqu'elle satisfait aux conditions du critère de Leibniz : 1) $1 > \frac{1}{2} > \frac{1}{3} > \ldots$; 2) $\lim \frac{1}{n} \to 0$ lorsque $n \to \infty$. Cette série s'appelle *série harmonique alternée.*

§ 4. Convergence absolue et semi-convergence des séries

Considérons maintenant les séries à termes réels. On appelle ainsi les séries composées de termes positifs et négatifs se suivant dans un ordre quelconque.

Soit une série à termes réels

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots = \sum_{n=1}^{\infty} a_n \tag{1}$$

et soit la série des modules des termes de (1)

$$|a_1| + |a_2| + |a_3| + \ldots + |a_n| + \ldots = \sum_{n=1}^{\infty} |a_n|. \tag{2}$$

On a le critère de convergence suivant pour les séries à termes réels.

Théorème 14.10. *Si la série* (2) *est convergente, la série* (1) *est convergente aussi et elle est dite absolument convergente.*

Démonstration. Supposons que la série (2) converge. Désignons par S_n la n-ième somme partielle de la série (1) et par σ_n celle de la série (2). La série (2) étant convergente, la suite de ses sommes partielles $\{\sigma_n\}$ converge vers σ et pour tout n on a

$$\sigma_n \leqslant \sigma, \tag{3}$$

puisque les termes de la série (2) sont strictement positifs.

Désignons par S_n' la somme des termes positifs et par S_n'' celle des modules des termes négatifs de S_n. Alors

$$S_n = S_n' - S_n'', \tag{4}$$

$$\sigma_n = S_n' + S_n''. \tag{5}$$

Il est évident que les suites $\{S_n'\}$ et $\{S_n''\}$ sont croissantes et de l'égalité (5) et l'inégalité (3) il s'ensuit qu'elles sont majorées : $S_n' \leqslant \sigma_n \leqslant \sigma$ et $S_n'' \leqslant \sigma_n \leqslant \sigma$. Donc $\lim\limits_{n\to\infty} S_n' = S'$ et $\lim\limits_{n\to\infty} S_n'' = S''$. Mais alors en vertu de (4)

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} (S_n' - S_n'') = \lim_{n\to\infty} S_n' - \lim_{n\to\infty} S_n'' = S' - S''.$$

Ce qui exprime que la série (1) converge. ■

Exemple 1. La série $1 - \frac{1}{2^2} - \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} - \frac{1}{6^2} - \frac{1}{7^2} + \ldots$ converge en vertu du critère précédent, puisque la série des modules correspondante converge (cf. exemple 6, § 2).

Le critère de convergence des séries à termes réels est suffisant mais pas nécessaire, puisqu'il existe des séries à termes réels qui convergent tandis que les séries de modules correspondantes di-

vergent. Ces séries sont dites *semi-convergentes*. Exemple: la série $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n}$. En effet, cette série converge d'après le critère de Leibniz (cf. exemple du § 3), alors que la série des modules correspondante, qui n'est autre que la série harmonique, diverge.

Exemple 2. La série $1-\frac{1}{2}+\frac{1}{4}-\frac{1}{8}+\frac{1}{16}-\frac{1}{32}+\ldots$ est absolument convergente, puisque la série des modules est convergente. (Ces deux séries sont des progressions géométriques de raisons $-1/2$ et $1/2$.)

Exemple 3. La série $1-\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}-\frac{1}{\sqrt{4}}+\frac{1}{\sqrt{5}}-\ldots$ est semi-convergente, car elle est convergente d'après le critère de Leibniz tandis que la série des modules est divergente (cf. exemple 6, § 2).

Signalons que la classification des séries convergentes en séries absolument convergentes et semi-convergentes est essentielle. Le fait est que les séries absolument convergentes jouissent de propriétés importantes que n'ont pas les séries semi-convergentes. Par exemple, dans une série semi-convergente la somme n'est pas égale à la somme des termes positifs et des termes négatifs comme c'est le cas pour une série absolument convergente (cf. démonstration du théorème 14.10).

§ 5. Séries entières

1. Définition et remarques générales. Une série de la forme

$$a_0+a_1x+a_2x^2+a_3x^3+\ldots+a_nx^n+\ldots=\sum_{n=0}^{\infty} a_nx^n \qquad (1)$$

s'appelle *série entière*.

Les nombres $a_0, a_1, a_2, a_3, \ldots, a_n, \ldots$ s'appellent *coefficients de la série entière*.

En attribuant à x des valeurs différentes, on obtient des séries numériques différentes qui peuvent être convergentes ou divergentes. L'ensemble des valeurs de x pour lesquelles la série (1) est convergente s'appelle *domaine de convergence* de la série (1). Cet ensemble n'est jamais vide, puisque toute série entière converge pour $x = 0$.

Il est évident que la somme partielle $S_n(x) = a_0 + a_1x + \ldots + a_nx^n$ est une fonction de la variable x. Donc la somme S de la série sera aussi une fonction de x définie sur le domaine de convergence de la série:

$$S=S(x)=\sum_{n=0}^{\infty} a_nx^n \quad \left(\text{ou } f(x)=\sum_{n=0}^{\infty} a_nx^n\right).$$

2. Intervalle de convergence d'une série entière. Prouvons un important théorème de la théorie des séries entières concernant le domaine de convergence d'une série entière.

Théorème 14.11 (Abel) *). 1) *Si une série entière* (1) *converge pour* $x = x_0$ ($x_0 \neq 0$), *elle converge absolument pour tout* x *satisfaisant à la condition* $|x| < |x_0|$; 2) *si une série* (1) *diverge pour* $x = x_1$, *elle divergera pour tout* x *tel que* $|x| > |x_1|$.

Démonstration. 1) La série numérique $\sum_{n=0}^{\infty} a_n x_0^n$ étant convergente par hypothèse, son terme général $a_n x_0^n \to 0$ lorsque $n \to \infty$, d'où il résulte que la suite $\{a_n x_0^n\}$ est majorée, c'est-à-dire qu'il existe un nombre $M > 0$ tel que

$$|a_n x_0^n| < M, \quad n = 0, 1, 2, \ldots \tag{2}$$

Mettons la série (1) sous la forme

$$a_0 + a_1 x_0 \left(\frac{x}{x_0}\right) + a_2 x_0^2 \left(\frac{x}{x_0}\right)^2 + \ldots + a_n x_0^n \left(\frac{x}{x_0}\right)^n + \ldots \tag{3}$$

et considérons la série des modules correspondante

$$|a_0| + |a_1 x_0| \left|\frac{x}{x_0}\right| + |a_2 x_0^2| \left|\frac{x}{x_0}\right|^2 + \ldots + |a_n x_0^n| \left|\frac{x}{x_0}\right|^n + \ldots \tag{4}$$

En vertu de (2) les termes de la série (4) sont strictement inférieurs aux termes correspondants de la série

$$M + M\left|\frac{x}{x_0}\right| + M\left|\frac{x}{x_0}\right|^2 + \ldots + M\left|\frac{x}{x_0}\right|^n + \ldots \tag{5}$$

Pour $|x| < |x_0|$ la série (5) est une progression géométrique de raison $q = \left|\frac{x}{x_0}\right| < 1$, donc elle est convergente. D'après le critère de comparaison la série (4) converge aussi, ce qui exprime que la série (1) converge absolument pour $|x| < |x_0|$.

2) Prouvons maintenant la deuxième partie du théorème. La série (1) diverge en x_1 par hypothèse. On demande de prouver qu'elle diverge pour tout x tel que $|x| > |x_1|$. Supposons par absurde qu'elle converge pour un x tel que $|x| > |x_1|$. D'après la première partie du théorème, la série (1) doit converger en x_1, puisque $|x_1| < |x|$, or ceci contredit le fait que la série diverge en x_1. ■

Le théorème d'Abel affirme que si une série entière converge en un point x_0, elle convergera absolument en tout point de l'in-

*) Abel Niels Henrik (1802-1829), mathématicien norvégien.

tervalle $]-|x_0|, |x_0|[$ (fig. 215, *a*); si elle diverge en x_1, elle divergera en tout point extérieur à l'intervalle $]-|x_1|, |x_1|[$ (fig. 215, *b*).

Ceci entraîne le

Théorème 14.12. *Si une série* $\sum_{n=0}^{\infty} a_n x^n$ *converge pour des* x *autres que* $x = 0$, *il existe un nombre* $R > 0$ *tel qu'elle converge absolument pour* $|x| < R$ *et diverge pour* $|x| > R$.

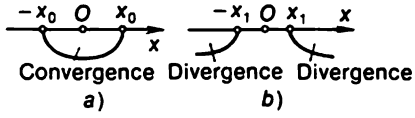


Fig. 215

Démonstration. Désignons par X le domaine de convergence de la série envisagée. Montrons que X est borné. En effet, si l'on considère un point x_1 en lequel la série diverge (de tels points existent, puisque la série converge pour certains x seulement), le théorème d'Abel nous dit que $|x| < |x_1|$ pour tout x de X. On sait que tout ensemble majoré admet une borne supérieure. Posons $R = \sup_{x \in X} |x|$. Cette série convergeant pour des x autres que $x = 0$, on a $R > 0$.

Considérons maintenant un x arbitraire tel que $|x| < R$. D'après la propriété de la borne supérieure, on peut exhiber un x_0 de X tel que $|x| < |x_0| \leqslant R$, ce qui, en vertu du théorème d'Abel, entraîne la convergence absolue pour la valeur de x envisagée.

Considérons maintenant un x arbitraire tel que $|x| > R$. Cet $x \notin X$. Donc la série diverge pour cet x. ■

L'intervalle $]-R, R[$ s'appelle *intervalle de convergence* de la série entière, le nombre R, *rayon de convergence*.

Signalons que l'intervalle de convergence de certaines séries est la droite numérique tout entière (dans ce cas on écrit $R = \infty$). Cet intervalle se réduit parfois à un point ($R = 0$).

Ainsi, toute série entière possède un rayon de convergence R.

Une série peut être convergente ou divergente pour $x = \pm R$. Cette question doit être étudiée pour chaque série envisagée.

Exhibons une méthode de détermination du rayon de convergence d'une série entière.

Théorème 14.13. *Si* $\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| \neq 0$ *existe, le rayon de convergence de la série* $\sum_{n=0}^{\infty} a_n x^n$ *est égal à* $R = \lim_{n \to \infty} \left| \frac{a_n}{a_{n+1}} \right|$.

Démonstration. Soit la série $\sum_{n=0}^{\infty} |a_n x^n|$. Posons $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{R}$ (cette limite existe et diffère de 0 par hypothèse). Alors

$$\lim_{n\to\infty} \left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right| = |x| \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \frac{|x|}{R}.$$

La série entière se transforme en série numérique pour chaque valeur de x. Donc d'après le critère de d'Alembert la série $\sum_{n=0}^{\infty} |a_n x^n|$ converge si $\frac{|x|}{R} < 1$, c'est-à-dire si $|x| < R$.

Donc la série $\sum_{n=0}^{\infty} a_n x^n$ converge absolument pour $|x| < R$ d'après le théorème 14.10, et diverge pour $|x| > R$, puisque $\lim_{n\to\infty} \left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right| = \frac{|x|}{R} > 1$. Son terme général $a_n x^n$ ne tend pas par conséquent vers 0 lorsque $n \to \infty$.

Donc la série envisagée est convergente à l'intérieur de l'intervalle $]-R, R[$ et divergente à l'extérieur, autrement dit son rayon de convergence est $R = \lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right|$. ■

Remarque. On démontre que si $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 0$, la série $\sum_{n=0}^{\infty} a_n x^n$ converge sur la droite numérique tout entière, c'est-à-dire que $R = \infty$, et si $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \infty$, elle converge seulement pour $x = 0$, c'est-à-dire que $R = 0$.

Exemple 1. Considérons la série $\sum_{n=1}^{\infty} \frac{|x^n|}{n}$. On a

$$R = \lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \lim_{n\to\infty} \frac{n+1}{n} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = 1.$$

Donc cette série est convergente sur l'intervalle $]-1, 1[$ d'après le théorème 14.13. Etudions cette série aux bornes ± 1 de l'intervalle de convergence. Pour $x = 1$, on obtient la série harmonique $\sum_{n=1}^{\infty} \frac{1}{n}$ qui, on le sait, est divergente, et pour $x = -1$, la série $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ qui est convergente d'après le

critère de Leibniz. Donc la série envisagée est convergente en tout point de l'intervalle $[-1, 1[$ et divergente à l'extérieur de cet intervalle.

Exemple 2. La série $\sum_{n=1}^{\infty} n!\, x^n$ est divergente sur toute la droite numérique sauf en $x=0$, puisque son rayon de convergence

$$R = \lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \lim_{n\to\infty} \frac{n!}{(n+1)!} = \lim_{n\to\infty} \frac{1}{1+n} = 0.$$

Exemple 3. La série $\sum_{n=1}^{\infty} \frac{x^n}{n!}$ converge absolument sur la droite numérique tout entière, puisque

$$R = \lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \lim_{n\to\infty} \frac{(n+1)!}{n!} = \lim_{n\to\infty} (n+1) = \infty.$$

3. Propriétés des séries entières. Supposons qu'une fonction $f(x)$ est la somme d'une série entière

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots, \qquad (6)$$

dont l'intervalle de convergence est $]-R, R[$.

On dit alors que la fonction $f(x)$ se développe en série entière de x sur l'intervalle $]-R, R[$.

Voici deux propriétés des séries entières que nous donnerons sous forme de théorèmes sans les prouver.

Théorème 14.14. *Si une fonction $f(x)$ se développe en une série entière* (6) *sur l'intervalle $]-R, R[$, elle est dérivable sur cet intervalle et sa dérivée $f'(x)$ peut être obtenue par une dérivation terme à terme de la série* (6), *c'est-à-dire que*

$$f'(x) = (a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots)' = \\ = a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1} + \ldots$$

Les dérivées de tout ordre de la fonction $f(x)$ se calculent par le même procédé et de plus les séries correspondantes ont le même intervalle de convergence que la série (6).

Théorème 14.15. *Si une fonction $f(x)$ se développe en série entière* (6) *sur l'intervalle $]-R, R[$, elle est intégrable sur cet intervalle et son intégrale peut être obtenue par une intégration terme à terme de la série* (6), *autrement dit, si $x_1, x_2 \in]-R, R[$,*

alors

$$\int_{x_1}^{x_2} f(x)\,dx = \int_{x_1}^{x_2} (a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots)\,dx =$$

$$= \int_{x_1}^{x_2} a_0\,dx + \int_{x_1}^{x_2} a_1x\,dx + \ldots + \int_{x_1}^{x_2} a_nx^n\,dx + \ldots$$

Il est intéressant d'intégrer la série entière (6) sur l'intervalle $[0, x]$, où $|x| < R$:

$$\int_0^x f(x)\,dx = a_0x + \frac{a_1x^2}{2} + \frac{a_2x^3}{3} + \ldots + \frac{a_nx^{n+1}}{n+1} + \ldots$$

Dans ce cas on obtient encore une série entière qui possède le même intervalle de convergence que la série (6).

On étudie parfois des séries entières de forme plus générale :

$$a_0 + a_1(x-a) + a_2(x-a)^2 + \ldots + a_n(x-a)^n + \ldots =$$

$$= \sum_{n=0}^{\infty} a_n(x-a)^n. \quad (7)$$

Les séries (7) se ramènent à la forme (1) par le changement $x - a = t$.

Si une fonction $f(x)$ est la somme d'une série (7), on dit qu'elle se développe en une série entière de $x - a$.

Tout ce qui a été dit sur les séries s'étend intégralement aux séries (7). Pour alléger l'écriture, on raisonnera uniquement sur des séries (1).

4. Développement des fonctions en séries entières. Le théorème suivant prouve qu'une fonction se développe d'une seule façon en série entière.

Théorème 14.16. *Si une fonction $f(x)$ se développe en une série entière*

$$f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots \quad (8)$$

sur l'intervalle $]-R, R[$, *ce développement est unique.*

Démonstration. Par hypothèse, la série (8) converge sur l'intervalle $]-R, R[$ et la fonction $f(x)$ est sa somme. Donc en vertu du théorème 14.4, la série (8) peut être dérivée terme à terme autant de fois qu'on veut sur l'intervalle $]-R, R[$. On

a alors

$$f'(x) = 1 \cdot a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + \ldots + na_nx^{n-1} + \ldots,$$
$$f''(x) = 1 \cdot 2a_2 + 2 \cdot 3a_3x + 3 \cdot 4a_4x^2 + 4 \cdot 5a_5x^3 + \ldots$$
$$\ldots + (n-1)\, na_nx^{n-2} + \ldots,$$
$$f'''(x) = 1 \cdot 2 \cdot 3a_3 + 2 \cdot 3 \cdot 4a_4x + \ldots + (n-2)\,(n-1)\, na_nx^{n-3} + \ldots,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$f^{(n)}(x) = n!\, a_n + 2 \cdot 3 \cdot 4 \ldots (n+1)\, a_{n+1}x + \ldots,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

En faisant $x = 0$ dans ces égalités et dans (8), on obtient

$$f(0) = a_0, \quad f'(0) = 1 \cdot a_1, \quad f''(0) = 2!\, a_2,$$
$$f'''(0) = 3!\, a_3, \ \ldots, \ f^{(n)}(0) = n!\, a_n, \ \ldots,$$

d'où

$$a_0 = f(0), \quad a_1 = \frac{f'(0)}{1!}, \quad a_2 = \frac{f''(0)}{2!}, \quad a_3 = \frac{f'''(0)}{3!}, \ \ldots$$
$$\ldots, \ a_n = \frac{f^{(n)}(0)}{n!}, \ \ldots \tag{9}$$

Ce qui prouve le théorème puisque les coefficients de la série (8) sont définis de façon unique par les formules (9). ■

En portant les expressions des coefficients dans (8), on obtient

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \ldots$$

Ainsi, si une fonction $f(x)$ se développe en série entière, cette série est de la forme

$$f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \ldots \tag{10}$$

La série (10) s'appelle *série de Maclaurin* de la fonction $f(x)$.

Supposons maintenant que $f(x)$ est une fonction indéfiniment dérivable. On peut la développer en une série (10). Voyons sous quelles conditions la somme de cette série (10) est confondue avec la fonction $f(x)$. La réponse à cette question nous est fournie par la formule de Maclaurin. Au chap. VI, § 3 on a montré que toute fonction indéfiniment dérivable était justiciable de la formule de Maclaurin

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + R_n(x),$$

où le reste

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}, \quad \xi = \theta x, \ 0 < \theta < 1. \tag{11}$$

Si l'on désigne par $S_n(x)$ la n-ième somme partielle de la série de Maclaurin, on peut écrire

$$f(x) = S_n(x) + R_n(x). \tag{12}$$

On a le théorème suivant.

Théorème 14.17. *Pour que la série de Maclaurin* (10) *soit convergente sur* $]-R, R[$ *et ait pour somme la fonction* $f(x)$, *il est nécessaire et suffisant que sur* $]-R, R[$ *le reste* $R_n(x)$ *de la formule de Maclaurin* (11) *tende vers* 0 *lorsque* $n \to \infty$, *c'est-à-dire que* $\lim\limits_{n\to\infty} R_n(x) = 0$ *pour tout* $x \in]-R, R[$.

Démonstration. *Nécessité.* Supposons que $f(x)$ est la somme de la série de Maclaurin sur $]-R, R[$, c'est-à-dire que $\lim\limits_{n\to\infty} S_n(x) = f(x)$. De l'égalité (12) il s'ensuit alors que $\lim\limits_{n\to\infty} R_n(x) = 0$ pour tout $x \in]-R, R[$.

Suffisance. Soit $\lim\limits_{n\to\infty} R_n(x) = 0$ pour tout $x \in]-R, R[$. De l'égalité (12) il s'ensuit alors que $\lim\limits_{n\to\infty} [f(x) - S_n(x)] = 0$, c'est-à-dire que $\lim\limits_{n\to\infty} S_n(x) = f(x)$. Ce qui exprime que la série de Maclaurin (10) est convergente sur $]-R, R[$ et sa somme égale à $f(x)$. ■

Ce théorème ramène donc le problème du développement d'une fonction en série de Maclaurin à l'étude du comportement du reste $R_n(x)$ lorsque $n \to \infty$.

Considérons le développement en série de Maclaurin de quelques fonctions élémentaires.

Développement de la fonction $f(x) = e^x$. On a $f'(x) = f''(x) = \ldots = f^{(n)}(x) = e^x$. Pour $x = 0$: $f(0) = f'(0) = f''(0) = \ldots = f^{(n)}(0) = 1$. La série de Maclaurin de e^x est donc

$$1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots \tag{13}$$

Trouvons l'intervalle de convergence de la série (13):

$$R = \lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \lim_{n\to\infty} \frac{n!\,(n+1)}{n!} = \infty.$$

Donc cette série est absolument convergente sur la droite tout entière.

Prouvons maintenant que la fonction e^x est la somme de la série (13).

D'après la condition nécessaire de convergence d'une série, on a pour tout x

$$\lim_{n\to\infty} \frac{|x|^n}{n!} = 0. \tag{14}$$

Comme $f^{(n+1)}(\xi) = e^{\xi}$, il vient

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\, x^{n+1} = \frac{e^{\xi}}{(n+1)!}\, x^{n+1},$$

où $\xi = \theta x$, $0 < \theta < 1$. D'où, compte tenu de ce que $e^{\xi} < e^{|x|}$,

$$|R_n(x)| = \frac{e^{\xi}}{(n+1)!}\, |x|^{n+1} < \frac{e^{|x|}}{(n+1)!}\, |x|^{n+1}.$$

Comme

$$\lim_{n\to\infty} \frac{|x|^n}{n!} = 0,$$

on a *a fortiori*

$$\lim_{n\to\infty} \frac{|x|^{n+1}}{(n+1)!} = 0.$$

En faisant tendre n vers ∞ dans la dernière inégalité, on trouve que $\lim\limits_{n\to\infty} R_n(x) = 0$ pour tout x, ce qui exprime que la fonction e^x est la somme de la série (13).

Donc pour tout x

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots$$

Développement de la fonction $f(x) = \sin x$. On a $f'(x) = \cos x = \sin\left(x + \frac{\pi}{2}\right)$, $f''(x) = -\sin x = \sin\left(x + 2\frac{\pi}{2}\right)$, $\ldots$, $f^{(n)}(x) = \sin\left(x + n\frac{\pi}{2}\right)$ (cf. chap. V, § 10, n° 2). Pour $x = 0$: $f(0) = 0$, $f'(0) = 1$, $f''(0) = 0$, $f'''(0) = -1$, $f^{(4)}(0) = 0$, $\ldots$ La série de Maclaurin de la fonction $\sin x$ est donc

$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + \frac{(-1)^{n-1} x^{2n-1}}{(2n-1)!} + \ldots$$

On s'assure immédiatement que cette série est absolument convergente sur la droite numérique tout entière. Etudions le reste

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\, x^{n+1} = \frac{\sin\left[\xi + (n+1)\frac{\pi}{2}\right]}{(n+1)!}\, x^{n+1},$$

où $\xi = \theta x$, $0 < \theta < 1$. Comme $\left|\sin\left[\xi + (n+1)\frac{\pi}{2}\right]\right| \leqslant 1$, il vient $|R_n(x)| \leqslant \frac{|x|^{n+1}}{(n+1)!}$. D'après (14), $\lim\limits_{n\to\infty} \frac{|x|^{n+1}}{(n+1)!} = 0$. Donc $\lim\limits_{n\to\infty} R_n(x) = 0$ pour tout x. Ce qui exprime que la fonction $\sin x$ est la somme de cette série, c'est-à-dire qu'on a le déve-

loppement

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + \frac{(-1)^{n-1} x^{2n-1}}{(2n-1)!} + \ldots$$

Développement de la fonction $f(x) = \cos x$. On peut obtenir le développement de $\cos x$ en série de Maclaurin en procédant comme ci-dessus. Mais il est simple de développer $\cos x$ en série en dérivant terme à terme la série de $\sin x$:

$$(\sin x)' = (x)' - \left(\frac{x^3}{3!}\right)' + \left(\frac{x^5}{5!}\right)' - \left(\frac{x^7}{7!}\right)' + \ldots$$

$$\ldots + \left[\frac{(-1)^{n-1} x^{2n-1}}{(2n-1)!}\right]' + \ldots,$$

d'où

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + \ldots$$

De nombreuses autres fonctions sont développables en série de Maclaurin. On aurait pu envisager aussi un développement plus général qu'en série de Maclaurin, par exemple un développement de Taylor suivant les puissances de $x - a$, où $a \neq 0$, de la forme

$$f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \ldots$$

Tout ce qui a été dit s'étend intégralement à ces séries.

En développant la fonction $\cos x$ en série de Maclaurin, nous nous sommes servi de la dérivabilité terme à terme des séries entières. On peut tout aussi bien user d'une autre propriété de ces séries : leur intégrabilité terme à terme. A titre d'exemple, développons les fonctions $\ln(1 + x)$ et $\operatorname{Arc\,tg} x$ en séries entières par une intégration terme à terme.

Soit la série $1 + x + x^2 + x^3 + \ldots + x^n + \ldots$ On reconnaît une progression géométrique de raison $q = x$. On sait que cette série est convergente pour $|x| < 1$ et sa somme est égale à $\frac{1}{1-x}$. Donc

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots + x^n + \ldots \qquad (15)$$

L'expression (15) est le développement de la fonction $f(x) = \frac{1}{1-x}$ en série entière.

En substituant $-t$ à x dans (15), on obtient

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \ldots + (-1)^n t^n + \ldots$$

pour $|t|<1$. Intégrons cette série terme à terme entre 0 et x ($|x|<1$). Nous obtenons

$$\int_0^x \frac{dt}{1+t} = \ln(1+t)\Big|_0^x = \ln(1+x) =$$

$$= \int_0^x (1 - t + t^2 - t^3 + \ldots + (-1)^n t^n + \ldots)\, dt =$$

$$= \int_0^x dt - \int_0^x t\, dt + \int_0^x t^2\, dt - \int_0^x t^3\, dt + \ldots + (-1)^n \int_0^x t^n\, dt + \ldots =$$

$$= t\Big|_0^x - \frac{t^2}{2}\Big|_0^x + \frac{t^3}{3}\Big|_0^x - \frac{t^4}{4}\Big|_0^x + \ldots + (-1)^n \frac{t^{n+1}}{n+1}\Big|_0^x + \ldots =$$

$$= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^n \frac{x^{n+1}}{n+1} + \ldots$$

D'où

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^n \frac{x^{n+1}}{n+1} + \ldots \quad (16)$$

L'expression (16) est le développement de la fonction $\ln(1+x)$ en série entière. Cette égalité est valable pour $|x|<1$. On démontre qu'elle l'est aussi pour $x = 1$.

En effet, pour $x = 1$ le premier membre est égal à $\ln 2$ et le second membre, à la série numérique

$$1 - \frac{1}{2} + \frac{1}{3} - \ldots + (-1)^{n-1} \frac{1}{n} + \ldots \quad (17)$$

qui est convergente d'après le critère de Leibniz. Reste à prouver que

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \ldots + (-1)^{n-1} \frac{1}{n} + \ldots \quad (18)$$

A cet effet, intégrons l'expression

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \ldots + (-1)^{n-1} t^{n-1} + (-1)^n \frac{t^n}{1+t}$$

entre 0 et 1. Nous obtenons

$$\int_0^1 \frac{dt}{1+t} = \ln(1+t)\Big|_0^1 = \ln 2 =$$

$$= \int_0^1 \left(1 - t + t^2 - t^3 + \ldots + (-1)^{n-1} t^{n-1} + (-1)^n \frac{t^n}{1+t}\right) dt =$$

$$= \int_0^1 dt - \int_0^1 t\, dt + \int_0^1 t^2\, dt - \int_0^1 t^3\, dt + \ldots + (-1)^{n-1} \int_0^1 t^{n-1}\, dt +$$

$$+(-1)^n \int_0^1 \frac{t^n}{1+t}\,dt = 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots+(-1)^{n-1}\frac{1}{n}+$$

$$+(-1)^n \int_0^1 \frac{t^n}{1+t}\,dt,$$

i.e.

$$\ln 2 = 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots+(-1)^{n-1}\frac{1}{n}+$$

$$+(-1)^n \int_0^1 \frac{t^n}{1+t}\,dt. \qquad (19)$$

Dans cette égalité la somme des n premiers termes est la somme partielle S_n de la série (17). Donc

$$\ln 2 - S_n = (-1)^n \int_0^1 \frac{t^n\,dt}{1+t}. \qquad (20)$$

Comme $\frac{t^n}{1+t} \leqslant t^n$ pour $0 \leqslant t \leqslant 1$, il vient $\left|(-1)^n \int_0^1 \frac{t^n\,dt}{1+t}\right| =$

$= \int_0^1 \frac{t^n\,dt}{1+t} \leqslant \int_0^1 t^n\,dt = \frac{t^{n+1}}{n+1}\Big|_0^1 = \frac{1}{n+1} \to 0$ lorsque $n \to \infty$. On en déduit que l'intégrale du second membre de (20) tend vers 0 lorsque $n \to \infty$ et par suite $\lim\limits_{n\to\infty} S_n = \ln 2$, ce qui prouve (18).

Trouvons maintenant le développement de Arc tg x. En substituant $-t^2$ à x dans (15) et en intégrant par rapport à t entre 0 et x, on obtient

$$\text{Arc tg } x = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + \ldots \qquad (21)$$

L'égalité (21) est valable pour $|x| < 1$. On pourrait démontrer comme plus haut qu'elle l'est aussi pour $x = \pm 1$.

Signalons en conclusion que les séries entières ont des applications variées. Elles permettent de calculer les valeurs des fonctions avec n'importe quelle précision (en particulier, les valeurs de π et e); de trouver les valeurs approchées des intégrales définies qui soit ne s'expriment pas par des fonctions élémentaires, soit sont difficiles à calculer. Ainsi, par exemple, l'intégrale $\int_0^a \frac{\sin x}{x}\,dx$ ne s'exprime pas par une fonction élémentaire, puisque la primitive de la fonction $\frac{\sin x}{x}$ n'est pas une fonction

élémentaire. Or cette primitive peut facilement être représentée par une série entière. En effet, puisque $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots$, en multipliant par $\frac{1}{x}$ on obtient

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \ldots,$$

cette série étant convergente pour tout x. En intégrant terme à terme entre 0 et a, on trouve

$$\int_0^a \frac{\sin x}{x}\,dx = a - \frac{a^3}{3!\,3} + \frac{a^5}{5!\,5} - \frac{a^7}{7!\,7} + \ldots$$

Cette égalité permet de calculer cette intégrale avec n'importe quelle précision pour tout a.

Enfin, les séries entières jouent un rôle primordial dans la résolution approchée des équations différentielles.

§ 6. Séries à termes complexes

1. Brefs rappels sur les nombres complexes. On appelle *nombre complexe* z un couple de réels (x, y) et on note $z = (x; y)$. x est la *partie réelle*, y la *partie imaginaire* de z.

Un nombre complexe $z = (x; y)$ se représente sur le plan Oxy par un point M de coordonnées $(x; y)$ ou par un vecteur $\overrightarrow{OM}$ dont les projections sur Ox et Oy sont respectivement égales à x et y (fig. 216). Le plan Oxy s'appelle alors *plan complexe*. M est l'*image* de z et z, l'*affixe* de M.

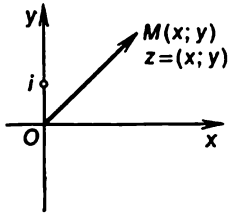


Fig. 216

Tout complexe de la forme $(x; 0)$ s'identifie au réel x, c'est-à-dire que $(x; 0) = x$. Ceci permet de traiter l'ensemble des réels comme un sous-ensemble des complexes. Sur le plan complexe, les réels sont figurés par les points de la droite Ox qui s'appelle *droite réelle*.

Les nombres complexes $(x; y)$, $y \neq 0$, sont encore appelés *imaginaires*. Les nombres de la forme $(0; y)$ s'appellent *imaginaires purs* et se notent $z = iy$. Le nombre $(0; 1)$ est l'*unité imaginaire* et se désigne par la lettre $i = (0; 1)$. Les nombres imaginaires purs sont figurés sur le plan complexe par les points de l'axe Oy qui est appelé *axe imaginaire*.

On dit que deux nombres complexes $z_1 = (x_1; y_1)$ et $z_2 = (x_2; y_2)$ sont égaux si $x_1 = x_2$ et $y_1 = y_2$. Le nombre complexe $0 = (0; 0)$ s'appelle *zéro*.

Opérations sur les nombres complexes. Définissons les opérations algébriques sur les nombres complexes.

On appelle *somme* des nombres complexes $z_1 = (x_1; y_1)$ et $z_2 = (x_2; y_2)$ le nombre complexe

$$z = (x_1 + x_2; y_1 + y_2).$$

On appelle *produit* des nombres complexes $z_1 = (x_1; y_1)$ et $z_2 = (x_2; y_2)$ le nombre complexe

$$z = (x_1x_2 - y_1y_2; x_1y_2 + x_2y_1).$$

La soustraction de deux nombres complexes se définit comme l'opération inverse de l'addition, c'est-à-dire qu'un nombre z est la *différence* d'un nombre z_1 et d'un nombre z_2 si $z + z_2 = z_1$. De cette définition il résulte que le nombre complexe

$$z = (x_1 - x_2; y_1 - y_2)$$

est la différence des nombres complexes $z_1 = (x_1; y_1)$ et $z_2 = (x_2; y_2)$.

La division des nombres complexes se définit comme l'opération inverse de la multiplication, c'est-à-dire qu'un nombre z est le *quotient* de z_1 par $z_2 \neq 0$ si $z \cdot z_2 = z_1$. Il s'ensuit de cette définition que le nombre complexe

$$z = \left(\frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2}; \frac{x_2y_1 - x_1y_2}{x_2^2 + y_2^2}\right)$$

est le quotient des nombres complexes $z_1 = (x_1; y_1)$ et $z_2 = (x_2; y_2)$.

L'unité imaginaire $i = (0; 1)$ joue un rôle particulier dans les opérations sur les nombres complexes. En l'élevant au carré, on obtient, en vertu de la définition du produit, $(0; 1) \cdot (0; 1) = (-1; 0) = -1$, c'est-à-dire que $i^2 = -1$. Donc tout nombre complexe $z = (x; y)$ se représente sous la forme

$$z = (x; y) = (x; 0) + (0; y) = (x; 0) + (y; 0) \cdot (0; 1) = x + iy, \quad (1)$$

appelée *forme algébrique* des nombres complexes. On voit à la forme (1) que les nombres complexes sont justiciables des opérations algébriques usuelles.

Le nombre complexe $\bar{z} = (x; -y) = x - iy$ s'appelle *conjugué complexe* du nombre $z = (x; y) = x + iy$ et se représente sur le plan complexe par le point symétrique à z par rapport à l'axe Ox.

Exemple. Calculer le produit des nombres complexes $z = x + iy$ et $\bar{z} = x - iy$.

Solution. On a

$$z \cdot \bar{z} = (x + iy) \cdot (x - iy) = xx + ixy - iyx + yy = x^2 + y^2.$$

Forme trigonométrique d'un nombre complexe. Considérons sur le plan complexe Oxy un système de coordonnées polaires de pôle O et d'axe polaire le demi-axe positif Ox (fig. 217). Considérons le nombre complexe $z = x + iy$. Les formules $x = \rho \cos \varphi$ et $y = \rho \sin \varphi$ qui relient les coordonnées polaires et rectangulaires nous permettent d'obtenir la forme trigonométrique du nombre complexe $z = x + iy$:

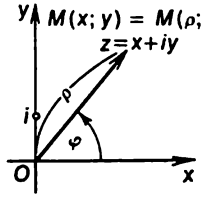


Fig. 217

$$z = \rho (\cos \varphi + i \sin \varphi).$$

Le nombre ρ s'appelle *module*, le nombre φ, *argument* de z. On les note $\rho = |z|$, $\varphi = \arg z$. L'argument φ est défini à $2k\pi$ près. Le module du nombre $z = x + iy$ vaut

$$\rho = |z| = |\overrightarrow{OM}| = \sqrt{x^2 + y^2}.$$

La valeur de l'argument φ telle que $\varphi \in [0, 2\pi[$ s'appelle *valeur principale* et se note $\varphi = \operatorname{Arg} z$. L'argument du nombre complexe $z = 0$ n'est pas défini et son module est nul.

Exemple. Représenter par leur forme trigonométrique les nombres complexes suivants: 1) $z = 2 + 2i\sqrt{3}$; 2) $z = 5$; 3) $z = i$.

Solution. 1) On a $x = 2, y = 2\sqrt{3}, \rho = \sqrt{(2)^2 + (2\sqrt{3})^2} = 4$, $\cos \varphi = x/\rho = 1/2$, $\sin \varphi = y/\rho = \sqrt{3}/2$. D'où $\varphi = \pi/3$. Donc $z = 2 + 2i\sqrt{3} = 4\left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3}\right)$.

2) $x = 5$, $y = 0$, $\rho = 5$, $\cos \varphi = 1$, $\sin \varphi = 0$. Donc $z = 5 = 5 (\cos 0 + i \sin 0)$.

3) $x = 0$, $y = 1$, $\rho = 1$, $\cos \varphi = 0$, $\sin \varphi = 1$. Donc $z = i = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2}$.

La forme trigonométrique nous permet de multiplier et de diviser facilement les nombres complexes. Soient $z_1 = \rho_1 (\cos \varphi_1 + i \sin \varphi_1)$ et $z_2 = \rho_2 (\cos \varphi_2 + i \sin \varphi_2)$. Alors

$$\begin{aligned} z = \rho (\cos \varphi + i \sin \varphi) &= z_1 \cdot z_2 = \\ &= \rho_1 (\cos \varphi_1 + i \sin \varphi_1) \cdot \rho_2 (\cos \varphi_2 + i \sin \varphi_2) = \\ &= \rho_1 \rho_2 [\cos \varphi_1 \cos \varphi_2 + i \cos \varphi_1 \sin \varphi_2 + \\ &\quad + i \sin \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2] = \\ &= \rho_1 \rho_2 [\cos (\varphi_1 + \varphi_2) + i \sin (\varphi_1 + \varphi_2)]. \end{aligned}$$

Donc

$$\rho = \rho_1 \rho_2, \quad \varphi = \varphi_1 + \varphi_2,$$

c'est-à-dire que la multiplication de deux nombres complexes conduit à multiplier leurs modules et à ajouter leurs arguments. La division de ces deux nombres complexes ($\rho_2 \neq 0$) nous conduit aux relations

$$\rho = \frac{\rho_1}{\rho_2}, \quad \varphi = \varphi_1 - \varphi_2.$$

Si l'on élève un nombre complexe $z = \rho(\cos\varphi + i\sin\varphi)$ à la puissance n, on obtient $z^n = \rho^n(\cos n\varphi + i\sin n\varphi)$. Si l'on pose $\rho = 1$, on obtient la formule

$$(\cos\varphi + i\sin\varphi)^n = \cos n\varphi + i\sin n\varphi,$$

appelée *formule de Moivre* *).

2. Limite d'une suite de nombres complexes. La théorie des limites développée pour les suites de nombres réels se généralise aux suites de nombres complexes; de nombreuses définitions relatives au passage à la limite se répètent intégralement.

Soit une suite de nombres complexes $z_1, z_2, \ldots, z_n, \ldots$ où $z_n = (x_n + iy_n)$ $(n = 1, 2, \ldots)$.

On dit qu'un nombre complexe $a = x_0 + iy_0$ est la *limite d'une suite* $\{z_n\}$ si pour tout $\varepsilon > 0$ il existe un entier N tel que $|z_n - a| < \varepsilon$ pour $n > N$. La suite $\{z_n\}$ est dite alors *convergente vers* a. Ceci se note $\lim\limits_{n\to\infty} z_n = a$ ou $z_n \to a$ lorsque $n \to \infty$.

Géométriquement cela signifie que pour tout $\varepsilon > 0$ tous les termes de $\{z_n\}$, à partir d'un certain rang N dépendant de ε, sont contenus dans un disque de rayon ε centré en a et les points z_n se rapprochent indéfiniment de a lorsque n croît. Ceci résulte des égalités $|z_n - a| = |(x_n - x_0) + i(y_n - y_0)| = \sqrt{(x_n - x_0)^2 + (y_n - y_0)^2} = \rho(M_n; M_0)$, où $\rho(M_n; M_0)$ est la distance des points $M_n(x_n; y_n)$ et $M_0(x_0; y_0)$, c'est-à-dire des images de z_n et a.

Tout nombre complexe z_n étant un couple de réels $(x_n; y_n)$, à toute suite $\{z_n\}$ correspondent deux suites de réels, $\{x_n\}$ et $\{y_n\}$, composées respectivement des parties réelles et imaginaires des termes z_n de la suite $\{z_n\}$.

On a le théorème suivant.

Théorème 14.18. *Une condition nécessaire et suffisante de convergence de la suite* $\{z_n\} = \{x_n + iy_n\}$ *est la convergence des suites de réels* $\{x_n\}$ *et* $\{y_n\}$.

Démonstration. *Nécessité.* Soit $z_n \to a = x_0 + iy_0$. Alors pour tout $\varepsilon > 0$ il existe un entier N tel que pour tout $n > N$ l'on a $|z_n - a| = \sqrt{(x_n - x_0)^2 + (y_n - y_0)^2} < \varepsilon$.

*) Moivre Abraham (de) (1667-1754), mathématicien anglais d'origine française.

D'où $|x_n - x_0| \leqslant |z_n - a| < \varepsilon$, $|y_n - y_0| \leqslant |z_n - a| < \varepsilon$ pour $n > N$ et par suite $x_n \to x_0$ et $y_n \to y_0$.

Suffisance. Soient $x_n \to x_0$, $y_n \to y_0$ et $a = x_0 + iy_0$. De la relation $|z_n - a| = \sqrt{(x_n - x_0)^2 + (y_n - y_0)^2}$ il s'ensuit alors que $z_n \to a$ lorsque $n \to \infty$. ■

Ce théorème permet de généraliser les principaux résultats de la théorie des limites de suites de nombres réels aux suites de nombres complexes.

3. Séries à termes complexes. Soit la série à termes complexes

$$z_1 + z_2 + \ldots + z_n + \ldots = \sum_{n=1}^{\infty} z_n, \tag{2}$$

où $z_n = x_n + iy_n$. On dit que la série (2) est *convergente* si converge la suite $\{S_n\}$ de ses sommes partielles $S_n = z_1 + z_2 + \ldots + z_n = \sum_{k=1}^{n} x_k + i \sum_{k=1}^{n} y_k$ lorsque $n \to \infty$; la limite $S = \lim_{n \to \infty} S_n$ s'appelle *somme* de la série (2). A la série (2) à termes complexes correspondent deux séries $\sum_{n=1}^{\infty} x_n$ et $\sum_{n=1}^{\infty} y_n$ à termes réels. Comme pour les suites de nombres complexes (cf. théorème 14.18), on démontre qu'une condition nécessaire et suffisante de convergence de la série (2) est la convergence des séries $\sum_{n=1}^{\infty} x_n$ et $\sum_{n=1}^{\infty} y_n$. En outre, si $\sum_{n=1}^{\infty} x_n = S'$ et $\sum_{n=1}^{\infty} y_n = S''$, alors $S = \sum_{n=1}^{\infty} z_n = S' + iS''$.

Le théorème suivant est utile à l'étude de la convergence de la série (2).

Théorème 14.19. *Soit donnée la série à termes complexes*

$$z_1 + z_2 + \ldots + z_n + \ldots = \sum_{n=1}^{\infty} z_n. \tag{3}$$

Si la série des modules

$$|z_1| + |z_2| + \ldots + |z_n| + \ldots = \sum_{n=1}^{\infty} |z_n| \tag{4}$$

est convergente, il en est de même de la série (3).

Démonstration. Soit $z_n = x_n + iy_n$; alors $|z_n| = \sqrt{x_n^2 + y_n^2}$ et $|x_n| \leqslant \sqrt{x_n^2 + y_n^2}$, $|y_n| \leqslant \sqrt{x_n^2 + y_n^2}$. Donc d'après le critère de comparaison des séries à termes réels (cf. théorème 14.6), la convergence de la série (4) entraîne celle des

séries $\sum_{n=1}^{\infty} |x_n|$ et $\sum_{n=1}^{\infty} |y_n|$. Donc les séries $\sum_{n=1}^{\infty} x_n$ et $\sum_{n=1}^{\infty} y_n$ convergent absolument, ce qui est une condition suffisante de convergence de la série (3). ■

Si la série (4) converge, on dit que la série (3) *converge absolument.*

Le théorème prouvé permet d'appliquer aux séries à termes complexes tous les critères de convergence des séries à termes positifs.

Exemple. Etudier la convergence de la série $\sum_{n=1}^{\infty} \frac{(1+i)^n}{n!}$.

Solution. Cette série converge absolument, puisque si l'on applique le critère de d'Alembert à la série des modules correspondante, on obtient

$$\lim_{n\to\infty} \frac{|z_{n+1}|}{|z_n|} = \lim_{n\to\infty} \frac{|1+i|^{n+1} \cdot n!}{(n+1)!\,|1+i|^n} = \lim_{n\to\infty} \frac{|1+i|}{n+1} =$$

$$= \lim_{n\to\infty} \frac{\sqrt{1^2+1^2}}{n+1} = \lim_{n\to\infty} \frac{\sqrt{2}}{n+1} = 0 < 1.$$

4. Séries entières à termes complexes. Une série de la forme

$$a_0 + a_1(z-a) + a_2(z-a)^2 + \ldots + a_n(z-a)^n + \ldots = \sum_{n=0}^{\infty} a_n(z-a)^n, \quad (5)$$

où z est une variable complexe et a_n et a, des nombres complexes, s'appelle *série entière.*

Comme pour les réels on se bornera à l'étude des séries entières de la forme

$$a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n + \ldots = \sum_{n=0}^{\infty} a_n z^n. \quad (6)$$

La série (5) se ramène à la forme (6) par le changement $z - a = t$.

Pour déterminer le domaine de convergence de la série entière (6) on se servira du théorème d'Abel (cf. théorème 14.11) qui s'énonce et se prouve comme pour les réels. Formulons ce théorème.

Théorème 14.20. 1) *Si la série entière* (6) *est convergente pour* $z = z_0$ $(z_0 \neq 0)$, *elle est absolument convergente pour tout* z *tel que* $|z| < |z_0|$; 2) *si la série* (6) *est divergente pour* $z = z_1$, *elle l'est pour tout* z *tel que* $|z| > |z_1|$.

Voyons l'interprétation géométrique du théorème d'Abel. Si la série est convergente en un point z_0, elle sera absolument

convergente en tout point du disque de rayon z_0 et de centre O; si elle est divergente en un point z_1, elle le sera en tout point extérieur au disque de rayon z_1 centré en O (fig. 218).

Comme dans le cas réel, on peut établir la forme du domaine de convergence de la série (6).

Théorème 14.21. *Si la série* (6) *est convergente pour des z autres que* $z = 0$, *il existe un nombre* $R > 0$ *tel qu'elle converge absolument pour* $|z| < R$ *et diverge pour* $|z| > R$.

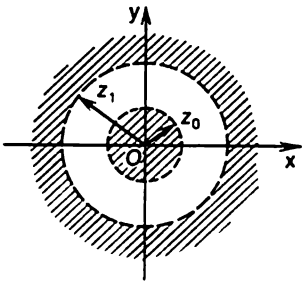


Fig. 218

Ce théorème se prouve comme le théorème 14.12. Le disque du plan complexe de centre O et de rayon R s'appelle *disque de convergence* de la série entière (6), et le nombre R, *rayon de convergence*. Comme dans le cas réel, la question de la convergence de la série sur la frontière du domaine de convergence (c'est-à-dire pour $|z| = R$) reste ouverte et implique une étude à part.

Si la série entière (6) ne converge qu'au point $z = 0$, on pose $R = 0$; si elle converge pour toute valeur de z, c'est-à-dire sur le plan complexe tout entier, alors on convient que $R = \infty$.

Le rayon de convergence de la série entière (6) se détermine par la même méthode que dans le cas réel.

Exemple. Trouver le rayon de convergence de la série entière $\sum_{n=1}^{\infty} \frac{z^n}{n!}$.

Solution. On a

$$R = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = \lim_{n\to\infty} \frac{(n+1)!}{n!} = \lim_{n\to\infty} (n+1) = \infty.$$

La série est donc convergente pour tout z.

Considérons la série

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n + \ldots = \sum_{n=0}^{\infty} a_n z^n. \qquad (7)$$

Il est évident que sa somme $f(z)$ est une fonction de la variable complexe z à l'intérieur du disque de convergence. Les fonctions d'une variable complexe qui se représentent par des séries entières sont dites *analytiques*. L'étude de ces fonctions sort du cadre de cet ouvrage.

5. Formule d'Euler. On se propose d'établir à l'aide des séries entières deux formules qui sont d'un vaste usage.

Au § 5 on a développé la fonction e^x en une série entière convergente pour tout x:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots \tag{8}$$

Si l'on substitue une variable complexe z à la variable x, on obtient la série entière de z:

$$1 + \frac{z}{1!} + \frac{z^2}{2!} + \ldots + \frac{z^n}{n!} + \ldots,$$

qui est convergente pour tout z en vertu de l'exemple étudié. Désignons sa somme par e^z. Donc par définition, pour tout nombre complexe z,

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \ldots + \frac{z^n}{n!} + \ldots \tag{9}$$

La somme de la série (9) s'appelle fonction exponentielle de la variable complexe z.

On définit de façon analogue les fonctions trigonométriques $\sin z$ et $\cos z$ de la variable complexe z:

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \ldots + (-1)^{n-1} \frac{z^{2n-1}}{(2n-1)!} + \ldots, \tag{10}$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \ldots + (-1)^n \frac{z^{2n}}{(2n)!} + \ldots \tag{11}$$

La fonction exponentielle e^z et les fonctions trigonométriques $\sin z$ et $\cos z$ sont reliées par une relation simple. Dans (9), substituons iz à z et regroupons tous les termes contenant i:

$$e^{iz} = 1 + iz + \frac{i^2z^2}{2!} + \frac{i^3z^3}{3!} + \frac{i^4z^4}{4!} + \frac{i^5z^5}{5!} + \ldots =$$

$$= 1 + iz - \frac{z^2}{2!} - i\frac{z^3}{3!} + \frac{z^4}{4!} + i\frac{z^5}{5!} - \ldots =$$

$$= \left(1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \ldots\right) + i\left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \ldots\right).$$

En comparant le résultat obtenu aux formules (10) et (11), on trouve

$$e^{iz} = \cos z + i \sin z. \tag{12}$$

En substituant $-iz$ à z dans (9), on obtient

$$e^{-iz} = \cos z - i \sin z. \tag{13}$$

Les formules (12) et (13) s'appellent *formules d'Euler*. Elles établissent un lien entre la fonction exponentielle et les fonctions trigonométriques $\sin z$ et $\cos z$. Si l'on ajoute et retranche terme

à terme les égalités (12) et (13), on obtient une autre forme d'écriture des formules d'Euler :

$$\cos z = \frac{e^{iz}+e^{-iz}}{2}; \qquad \sin z = \frac{e^{iz}-e^{-iz}}{2i}. \tag{14}$$

Les fonctions e^z, $\sin z$ et $\cos z$ sont donc reliées par les formules (12), (13) et (14). Pour $z = x$ (x réel) ces fonctions sont confondues respectivement avec les fonctions e^x, $\sin x$ et $\cos x$ de la variable réelle x.

§ 7. Séries de Fourier

1. La série trigonométrique et ses propriétés fondamentales. On appelle *série trigonométrique* une série de la forme

$$\frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \ldots + a_n \cos nx +$$

$$+ b_n \sin nx + \ldots = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{1}$$

où a_0, a_1, b_1, a_2, b_2, . . ., a_n, b_n, . . . sont les *coefficients* de cette série.

A la différence de la série entière qui fait intervenir les fonctions élémentaires $1, x, x^2, \ldots, x^n, \ldots$ la série trigonométrique est composée des fonctions circulaires

$$1/2, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx, \ldots, \tag{2}$$

qui sont bien étudiées aussi.

Signalons tout d'abord que les fonctions du système (2) sont toutes 2π-périodiques. En effet, la constante $1/2$ est de période quelconque, les fonctions $\sin nx$ et $\cos nx$ ($n = 1, 2, \ldots$) de période $2\pi/n$ *), donc aussi de période $2\pi = n(2\pi/n)$. Il est évident que chaque terme de la série trigonométrique (1) est une fonction 2π-périodique. Donc toute somme partielle de la série (1) est 2π-périodique (si les termes de la série ne changent pas lorsqu'on remplace x par $x + 2\pi$, il en est de même de leur somme). On en déduit que si la série (1) converge sur l'intervalle $[-\pi, \pi]$, elle converge sur la droite numérique tout entière et sa somme, qui est la limite d'une suite de sommes partielles périodiques, est une fonction 2π-périodique. Donc les séries trigonométriques sont particulièrement bien adaptées à l'étude des fonctions périodiques utilisées pour décrire les processus périodiques tels que les mouvements oscillatoires et circulaires des pièces d'une machine ou d'un engin, le mouvement périodique des corps célestes et des particules élémentaires, les oscillations acoustiques et électromagnétiques, etc.

*) En effet, $\sin [n(x \pm 2\pi/n)] = \sin (nx \pm 2\pi) = \sin nx$.

Une autre propriété importante des fonctions du système (2) est leur orthogonalité sur l'intervalle $[-\pi, \pi]$ au sens suivant : l'intégrale sur $[-\pi, \pi]$ du produit de deux fonctions différentes de ce système est nulle et l'intégrale sur $[-\pi, \pi]$ du carré de toute fonction de ce système est non nulle.

En effet,

$$\int_{-\pi}^{\pi} \frac{1}{2} \cos kx \, dx = \frac{1}{2k} \sin kx \Big|_{-\pi}^{\pi} = 0; \qquad (3)$$

$$\int_{-\pi}^{\pi} \frac{1}{2} \sin kx \, dx = -\frac{1}{2k} \cos kx \Big|_{-\pi}^{\pi} = 0.$$

Par ailleurs,

$$\int_{-\pi}^{\pi} \cos kx \cos nx \, dx = \frac{1}{2} \int_{-\pi}^{\pi} [\cos (k+n) x + \cos (k-n) x] \, dx =$$

$$= \frac{1}{2} \left[\frac{\sin (k+n) x}{k+n} + \frac{\sin (k-n) x}{k-n} \right]_{-\pi}^{\pi} = 0 \text{ pour } k \neq n. \qquad (4)$$

De façon analogue,

$$\int_{-\pi}^{\pi} \sin kx \sin nx \, dx = 0 \text{ pour } k \neq n; \quad \int_{-\pi}^{\pi} \sin kx \cos nx \, dx = 0. \qquad (5)$$

Enfin,

$$\left.\begin{aligned} \int_{-\pi}^{\pi} \cos^2 kx \, dx &= \frac{1}{2} \int_{-\pi}^{\pi} (1 + \cos 2kx) \, dx = \\ &= \frac{1}{2} \left[x + \frac{1}{2k} \sin 2kx \right]_{-\pi}^{\pi} = \pi, \\ \int_{-\pi}^{\pi} \sin^2 kx \, dx &= \frac{1}{2} \int_{-\pi}^{\pi} (1 - \cos 2kx) \, dx = \\ &= \frac{1}{2} \left[x - \frac{1}{2k} \sin 2kx \right]_{-\pi}^{\pi} = \pi, \\ \int_{-\pi}^{\pi} \left(\frac{1}{2} \right)^2 &= \frac{1}{4} x \Big|_{-\pi}^{\pi} = \frac{\pi}{2}, \end{aligned}\right\} \qquad (6)$$

C.Q.F.D.

2. Série de Fourier. De même que pour les séries entières, on a le théorème suivant pour les séries trigonométriques.

Théorème 14.22. *Si une fonction $f(x)$ définie et intégrable sur l'intervalle $[-\pi, \pi]$ se développe en une série trigonométrique*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{7}$$

intégrable terme à terme, alors ce développement est unique.

Démonstration. Une intégration de (7) nous donne

$$\int_{-\pi}^{\pi} f(x)\,dx = \frac{a_0}{2} \int_{-\pi}^{\pi} dx + \sum_{n=1}^{\infty} \left[a_n \int_{-\pi}^{\pi} \cos nx\,dx + b_n \int_{-\pi}^{\pi} \sin nx\,dx \right],$$

d'où l'on déduit, compte tenu de (3),

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\,dx. \tag{8}$$

Pour déterminer le coefficient a_k de $\cos kx$ multiplions (7) par $\cos kx$ et intégrons entre $-\pi$ et π *). Les formules (3) à (6) nous donnent

$$\int_{-\pi}^{\pi} f(x) \cos kx\,dx = \frac{a_0}{2} \int_{-\pi}^{\pi} \cos kx\,dx + \sum_{n=1}^{\infty} \left[a_n \int_{-\pi}^{\pi} \cos kx \cos nx\,dx + \right.$$

$$\left. + b_n \int_{-\pi}^{\pi} \cos kx \sin nx\,dx \right] = a_k \int_{-\pi}^{\pi} \cos^2 kx\,dx = a_k \pi,$$

d'où

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx\,dx. \tag{9}$$

De façon analogue, en multipliant (7) par $\sin kx$ et en intégrant entre $-\pi$ et π, on obtient en vertu des mêmes formules

$$\int_{-\pi}^{\pi} f(x) \sin kx\,dx = b_k \pi,$$

d'où

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx\,dx. \tag{10}$$

Donc les coefficients a_0, a_k et b_k de la série (7) sont définis de façon unique par les formules (8), (9) et (10), ce qui prouve le théorème. ■

*) On démontre que la série (7) est intégrable terme à terme après sa multiplication par une fonction bornée.

Ce théorème nous permet d'introduire la définition suivante.

Définition. Soit $f(x)$ une fonction définie et intégrable sur l'intervalle $[-\pi, \pi]$. Les nombres a_0, a_n et b_n déterminés à l'aide des formules (8), (9) et (10) s'appellent *coefficients de Fourier* et la série

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

série de Fourier *) de la fonction $f(x)$.

3. Convergence de la série de Fourier. Introduisons la notion de prolongement périodique d'une fonction $f(x)$ définie sur l'intervalle $[-\pi, \pi]$.

On dira qu'une fonction $F(x)$ 2π-périodique définie sur la droite numérique tout entière est le *prolongement périodique* d'une fonction $f(x)$ si $F(x) = f(x)$ sur l'intervalle $[-\pi, \pi]$.

Il est évident que si une série de Fourier converge sur l'intervalle $[-\pi, \pi]$ vers une fonction $f(x)$, elle convergera sur la droite numérique tout entière vers son prolongement périodique $F(x)$.

Voyons sous quelles conditions la série de Fourier d'une fonction $f(x)$ converge vers cette fonction.

Théorème 14.23. *Supposons qu'une fonction $f(x)$ et sa dérivée $f'(x)$ sont ou continues, ou affectées d'un nombre fini de discontinuités de première espèce sur l'intervalle $[-\pi, \pi]$. Alors la série de Fourier de $f(x)$ est convergente sur la droite numérique tout entière et en outre la somme de cette série est égale à $f(x)$ en tout point $x \in]-\pi, \pi[$ où $f(x)$ est continue et à*

$$\frac{f(x_0-)+f(x_0+)}{2},$$

où $f(x_0-) = \lim\limits_{x \to x_0-} f(x)$ et $f(x_0+) = \lim\limits_{x \to x_0+} f(x)$, en tout point x_0 où $f(x)$ est discontinue. Aux bornes de l'intervalle $[-\pi, \pi]$ la somme de la série est égale à

$$\frac{f(-\pi)+f(\pi)}{2}.$$

En tout point x extérieur à l'intervalle $[-\pi, \pi]$ la somme de la série est égale au prolongement périodique $F(x)$ de $f(x)$ si $F(x)$ est continue en x et à $\frac{F(x-)+F(x+)}{2}$ si $F(x)$ est discontinue en x.

---

*) Fourier Jean-Baptiste Joseph (1768-1830), mathématicien et physicien français.

4. Séries de Fourier de fonctions paires et impaires. Soit $f(x)$ une fonction paire, i.e. $f(x) = f(-x)$, définie sur l'intervalle $[-\pi, \pi]$. Ses coefficients de Fourier b_n sont alors nuls. En effet,

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx =$$

$$= \frac{1}{\pi} \left[\int_{-\pi}^{0} f(x) \sin nx \, dx + \int_{0}^{\pi} f(x) \sin nx \, dx \right].$$

Dans la première intégrale entre crochets faisons le changement de variable: $x = -t$. Alors $dx = -dt$; si $x = -\pi$, $t = \pi$; si $x = 0$, $t = 0$. Comme $f(x)$ est paire et $\sin x$ impair, on obtient

$$\int_{-\pi}^{0} f(x) \sin nx \, dx = -\int_{-\pi}^{0} f(-t) \sin n(-t) \, dt = -\int_{0}^{\pi} f(t) \sin nt \, dt.$$

Donc

$$b_n = \frac{1}{\pi} \left[-\int_{0}^{\pi} f(t) \sin nt \, dt + \int_{0}^{\pi} f(x) \sin nx \, dx \right] = 0$$

(on rappelle que l'intégrale définie ne dépend pas du nom de la variable d'intégration).

De façon analogue, en tenant compte de la parité des fonctions $f(x)$ et $\cos x$, on obtient les expressions suivantes pour les coefficients a_n :

$$a_0 = \frac{2}{\pi} \int_{0}^{\pi} f(x) \, dx, \quad a_n = \frac{2}{\pi} \int_{0}^{\pi} f(x) \cos nx \, dx. \tag{11}$$

Supposons maintenant que $f(x)$ est une fonction impaire, i.e. $f(x) = -f(-x)$, définie sur l'intervalle $[-\pi, \pi]$. En reprenant les mêmes raisonnements, on démontre que les coefficients de Fourier a_n sont nuls et les coefficients b_n, définis par

$$b_n = \frac{2}{\pi} \int_{0}^{\pi} f(x) \sin nx \, dx. \tag{12}$$

Donc la série de Fourier d'une fonction $f(x)$ est une série de cosinus si $f(x)$ est paire et une série de sinus si $f(x)$ est impaire. Les formules (11) et (12) nous permettent d'alléger le calcul des coefficients de Fourier lorsque la fonction $f(x)$ est paire ou impaire.

Exemple 1. Soit la fonction $f(x) = x$. Cette fonction satisfait aux conditions du théorème 14.23. Elle se développe donc en série de Fourier. Vu qu'elle est paire, ses coefficients a_n sont nuls

et ses coefficients b_n, définis par la formule (12). On a

$$b_n = \frac{2}{\pi}\int_0^{\pi} x \sin nx\,dx = \frac{2}{\pi}\left[-\frac{1}{n}x\cos nx\Big|_0^{\pi} + \right.$$

$$\left. + \frac{1}{n}\int_0^{\pi}\cos nx\,dx\right] = (-1)^{n+1}\frac{2}{n}.$$

La série de Fourier de x est donc

$$x - 2\left(\frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \right.$$

$$\left. - \frac{\sin 4x}{4} + \ldots + (-1)^{n+1}\frac{\sin nx}{n} + \ldots\right).$$

Cette égalité est vérifiée pour tout $x \in]-\pi, \pi[$. Aux points $x = \pm\pi$, la somme de la série de Fourier n'est pas confondue en

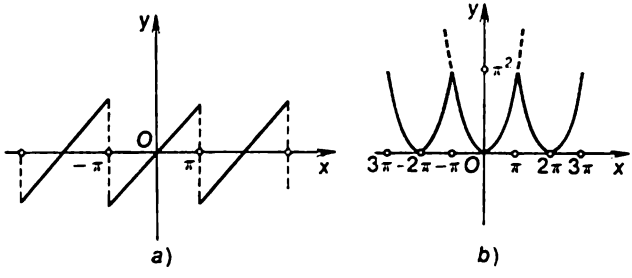


Fig. 219

vertu du théorème 14.23 avec les valeurs de la fonction $f(x) = x$, mais vaut $\frac{f(-\pi)+f(\pi)}{2} = \frac{-\pi+\pi}{2} = 0$. A l'extérieur de $[-\pi, \pi]$ la somme de la série est le prolongement périodique de la fonction $f(x) = x$; son graphique est représenté sur la figure 219, *a*.

Exemple 2. Considérons la fonction $f(x) = x^2$. Cette fonction satisfaisant à toutes les conditions du théorème 14.23 se développe en série de Fourier. Comme elle est paire, ses coefficients de Fourier $b_n = 0$ et ses coefficients a_n sont définis par les formules (11). On a

$$a_0 = \frac{2}{\pi}\int_0^{\pi} x^2\,dx = \frac{2\pi^2}{3};\quad a_n = \frac{2}{\pi}\int_0^{\pi} x^2\cos nx\,dx =$$

$$= \frac{2}{\pi}\left[\frac{x^2\sin nx}{n}\Big|_0^{\pi} - \frac{2}{n}\int_0^{\pi} x\sin nx\,dx\right] = (-1)^n\frac{4}{n^2}.$$

Donc la série de Fourier de x^2 est de la forme

$$x^2 = \frac{\pi^2}{3} - 4\left(\frac{\cos x}{1} - \frac{\cos 2x}{2^2} + \frac{\cos 3x}{3^2} - \ldots\right).$$

Cette égalité est valable pour tout $x \in [-\pi, \pi]$, puisque aux points $x = \pm\pi$ la somme de la série est confondue avec les valeurs de $f(x) = x^2$. En effet, $\frac{f(-\pi)+f(\pi)}{2} = \frac{\pi^2+\pi^2}{2} = \pi^2 = f(\pi) = f(-\pi)$. Les graphiques de la fonction $f(x) = x^2$ et de la somme de cette série de Fourier sont représentés sur la figure 219, *b*.

5. Série de Fourier de période $2l$.

Soit $f(x)$ une fonction définie et satisfaisant aux conditions du théorème 14.23 sur un intervalle $[-l, l]$ ($l > 0$ est un nombre arbitraire). Développons-la en série de Fourier.

Faisons le changement

$$x = \frac{l\xi}{\pi}$$

et considérons la fonction $\varphi(\xi) = f\left(\frac{l\xi}{\pi}\right) = f(x)$.

Il est évident que la fonction $\varphi(\xi)$ est définie et satisfait aux conditions du théorème 14.23 sur l'intervalle $[-\pi, \pi]$.

Développons la fonction $\varphi(\xi)$ en série de Fourier sur l'intervalle $[-\pi, \pi]$

$$\varphi(\xi) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\xi + b_n \sin n\xi), \qquad (13)$$

où

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} \varphi(\xi)\, d\xi; \quad a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \varphi(\xi)\cos n\xi\, d\xi;$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \varphi(\xi)\sin n\xi\, d\xi.$$

Revenons maintenant à l'ancienne variable x : $x = \frac{l}{\pi}\xi$, $\xi = x\frac{\pi}{l}$, $d\xi = \frac{\pi}{l}dx$. La formule (13) devient alors

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi x}{l} + b_n \sin\frac{n\pi x}{l}\right), \qquad (14)$$

où

$$a_0 = \frac{1}{l}\int_{-l}^{l} f(x)\,dx;\quad a_n = \frac{1}{l}\int_{-l}^{l} f(x)\cos\frac{n\pi x}{l}\,dx;$$

$$b_n = \frac{1}{l}\int_{-l}^{l} f(x)\sin\frac{n\pi x}{l}\,dx.$$

La formule (14) est une série de Fourier de période $2l$.

Exemple 3. Développer la fonction $f(x) = |x|$ en une série de Fourier de période $2l$ sur l'intervalle $[-l, l]$.

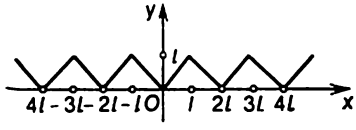


Fig. 220

Solution. La fonction $f(x) = |x|$ étant paire, on a $b_n = 0$;

$$a_0 = \frac{2}{l}\int_0^l x\,dx = \frac{2}{l}\cdot\frac{x^2}{2}\Big|_0^l = l;\quad a_n = \frac{2}{l}\int_0^l x\cos\frac{n\pi x}{l}\,dx =$$

$$= \frac{2}{l}\left[\frac{l\cdot x\sin\frac{n\pi x}{l}}{n\pi}\Big|_0^l - \frac{l}{n\pi}\int_0^l \sin\frac{n\pi x}{l}\,dx\right] = \frac{2l}{n^2\pi^2}\cos\frac{n\pi x}{l}\Big|_0^l =$$

$$= \frac{2l}{n^2\pi^2}[(-1)^n - 1] = \begin{cases} 0 & \text{pour } n \text{ pair,} \\ -\dfrac{4l}{n^2\pi^2} & \text{pour } n \text{ impair.} \end{cases}$$

Donc la série de Fourier de $f(x)$ est de la forme

$$|x| = \frac{l}{2} - \frac{4l}{\pi^2}\left[\cos\frac{\pi x}{l} + \frac{1}{3^2}\cos\frac{3\pi x}{l} + \frac{1}{5^2}\cos\frac{5\pi x}{l} + \ldots\right].$$

La fonction $|x|$ satisfait aux conditions du théorème 14.23 et cette égalité est valable pour tout $x \in [-l, l]$, ce qui exprime que la série converge sur la droite numérique tout entière et sa somme est la fonction dont le graphique est représenté sur la figure 220.

Signalons que les séries de Fourier sont largement utilisées, aussi bien en théorie qu'en pratique.

CHAPITRE XV

ÉQUATIONS DIFFÉRENTIELLES ORDINAIRES

Les équations différentielles occupent une place particulière en mathématiques. L'étude des phénomènes naturels les plus divers conduit souvent à la résolution de telles équations, car les lois qui régissent ces phénomènes sont décrites par des équations différentielles.

Les équations différentielles sont des équations dans lesquelles la fonction inconnue figure par sa ou ses dérivées. Le problème fondamental de la théorie des équations différentielles est l'étude des fonctions qui sont solutions de ces équations.

Les équations différentielles se divisent en équations différentielles ordinaires dans lesquelles les fonctions inconnues dépendent d'une seule variable et en équations différentielles aux dérivées partielles dans lesquelles les fonctions inconnues dépendent de deux ou de plusieurs variables.

La théorie des équations différentielles aux dérivées partielles est plus compliquée et fait l'objet d'ouvrages plus complets ou d'ouvrages spéciaux. Dans ce chapitre on expose les éléments de théorie des équations différentielles ordinaires. Dans la suite, par équations différentielles on n'entendra que des équations différentielles ordinaires.

§ 1. Equations différentielles du premier ordre

Commençons l'étude des équations différentielles par le cas le plus simple, celui d'une équation du premier ordre.

1. Définition d'une équation différentielle du premier ordre.

Définition 1. On appelle *équation différentielle du premier ordre* une équation de la forme

$$F(x, y, y') = 0, \tag{1}$$

où x est la variable indépendante; y la fonction inconnue; y' sa dérivée première.

Si l'équation (1) est résoluble par rapport à y', elle prend la forme

$$y' = f(x, y) \tag{2}$$

et s'appelle *équation du premier ordre résolue par rapport à la dérivée*. Etudions ces équations.

Dans certains cas il est plus avantageux d'écrire l'équation (2) sous la forme $\frac{dy}{dx} = f(x, y)$ ou sous la forme $f(x, y)\,dx - dy = 0$ qui est un cas particulier de l'équation plus générale

$$P(x, y)\,dx + Q(x, y)\,dy = 0, \tag{3}$$

où $P(x, y)$ et $Q(x, y)$ sont des fonctions connues. La représentation symétrique (3) est intéressante par le fait que les variables x et y jouent le même rôle, c'est-à-dire que chacune d'elles peut être traitée comme une fonction de l'autre.

Citons des exemples d'équations différentielles de la forme (2) et (3):

$$y' = xe^y, \quad y' = \frac{y \ln x}{x}, \quad y' = x + y, \quad x\,dx + y\,dy = 0.$$

2. Solution d'une équation. Problème de Cauchy.

Définition 2. On appelle *solution d'une équation différentielle du premier ordre* une fonction $y = \varphi(x)$, $x \in]a, b[$ *) qui transforme cette équation en identité.

Ainsi la fonction $y = x^3$, $x \in]-\infty, +\infty[$ est solution de l'équation $3y - xy' = 0$.

La courbe représentative d'une solution s'appelle *courbe intégrale*.

A la question de savoir sous quelles conditions l'équation (2) admet une solution répond le théorème de Cauchy qui est un théorème fondamental en théorie des équations différentielles, appelé théorème d'existence et d'unicité de la solution de l'équation différentielle (2).

Théorème 15.1 (Cauchy). *Si une fonction $f(x, y)$ et sa dérivée partielle $f'_y(x, y)$ sont définies et continues sur un domaine G du plan Oxy, l'équation $y' = f(x, y)$ admet dans un voisinage de tout point $(x_0; y_0)$ intérieur à G une seule solution satisfaisant aux conditions*

$$y = y_0 \quad pour \quad x = x_0. \tag{4}$$

Le théorème de Cauchy permet, à la forme de l'équation différentielle (2), de résoudre le problème d'existence et d'unicité de la solution de cette équation. Ceci est particulièrement important lorsqu'on ne sait pas à l'avance si cette équation possède une solution.

*) Cet intervalle peut être fini ou infini.

Géométriquement, le théorème affirme que par chaque point intérieur $(x_0; y_0)$ du domaine G il passe une seule courbe intégrale. Il est évident que l'équation (2) admet une infinité de solutions dans le domaine G.

Les conditions (4) en vertu desquelles une solution $y = \varphi(x)$ prend une valeur donnée y_0 en un point x_0 donné s'appellent *conditions initiales* de la solution et s'écrivent

$$y|_{x=x_0} = y_0. \tag{5}$$

La recherche de la solution de l'équation (2) qui vérifie des conditions initiales (5) est l'un des plus importants problèmes de théorie des équations différentielles. Ce problème s'appelle *problème de Cauchy*. Du point de vue géométrique, résoudre le problème de Cauchy c'est déterminer parmi l'ensemble infini de courbes intégrales celle qui passe par un point donné $(x_0; y_0)$ du plan Oxy.

Les points du plan par lesquels il passe plus d'une courbe intégrale ou aucune s'appellent *points singuliers* de l'équation.

3. Solutions générale et particulière d'une équation. Donnons deux définitions fondamentales.

Définition 3. Une fonction $y = \varphi(x, C)$ dépendant de x et d'une constante C s'appelle *solution générale* de l'équation (2) dans un domaine G du plan Oxy si elle est solution de l'équation (2) pour toute valeur de la constante C et si, quelles que soient les conditions initiales (5) telles que $(x_0; y_0) \in G$, il existe une seule valeur C_0 de C telle que la fonction $y = \varphi(x, C_0)$ satisfait aux conditions initiales $\varphi(x_0, C_0) = y_0$.

Définition 4. On appelle *solution particulière* de l'équation (2) dans le domaine G une solution $y = \varphi(x, C_0)$ déduite de la solution générale $y = \varphi(x, C)$ en donnant à C une valeur particulière C_0.

Géométriquement, la solution générale $y = \varphi(x, C)$ représente une famille de courbes intégrales du plan Oxy, dépendant d'une seule constante arbitraire C; la solution particulière $y = \varphi(x, C_0)$, la courbe intégrale de cette famille, qui passe par le point donné $(x_0; y_0)$.

Les conditions initiales (5) sont parfois appelées *conditions de Cauchy* et la solution particulière correspondante, *solution d'un problème de Cauchy*.

Exemple 1. Soit l'équation $y' = 3x^2$.

C'est une équation différentielle du premier ordre. Elle satisfait à toutes les conditions du théorème de Cauchy, puisque les fonctions $f(x, y) = 3x^2$ et $f'_y(x, y) = 0$ sont définies et continues sur le plan Oxy tout entier. Il est immédiat de vérifier que

la fonction $y = x^3 + C$, où C est une constante arbitraire, est la solution générale de cette équation sur le plan Oxy tout entier.

La solution générale représente une famille de paraboles cubiques. En donnant à C des valeurs différentes, on obtient des solutions différentes. Exemple: $y = x^3$ pour $C = 0$; $y = x^3 - 1$ pour $C = -1$, etc.

Posons un problème de Cauchy, c'est-à-dire cherchons une solution vérifiant les conditions initiales $x = x_0$ et $y = y_0$. En

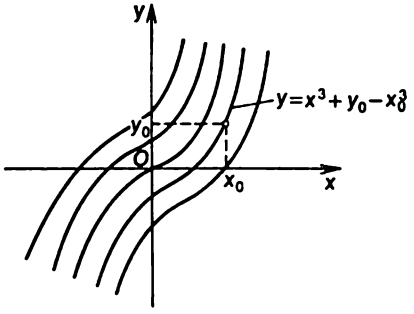


Fig. 221

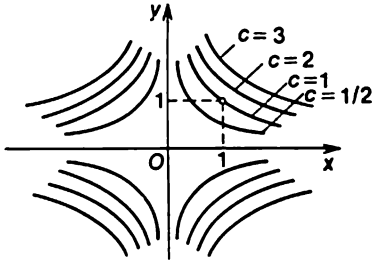


Fig. 222

portant ces valeurs dans la solution générale $y = x^3 + C$, on obtient $y_0 = x_0^3 + C$, d'où $C = y_0 - x_0^3$. Donc la solution particulière cherchée est $y = x^3 + y_0 - x_0^3$. Géométriquement, cela revient à choisir parmi les paraboles cubiques $y = x^3 + C$ celle qui passe par le point $(x_0; y_0)$ (fig. 221).

Exemple 2. Soit l'équation $y' = -y/x$.

C'est une équation différentielle du premier ordre. Les fonctions $f(x, y) = -y/x$ et $f'_y(x, y) = -1/x$ étant continues pour $x \neq 0$, cette équation satisfait aux conditions du théorème de Cauchy sur le plan Oxy privé de l'axe Oy.

Il est immédiat de vérifier que la solution générale de cette équation dans les domaines $y > 0$ et $y < 0$ est la fonction $y = C/x$, où C est une constante arbitraire. On obtient des solutions différentes en donnant à la constante C des valeurs différentes.

Trouvons une solution particulière, par exemple celle qui satisfait les conditions initiales $x_0 = 1$, $y_0 = 1$. On a $1 = C/1$. D'où $C = 1$ et la solution particulière cherchée est $y = 1/x$.

La solution générale de cette équation représente une famille d'hyperboles $y = C/x$. En se donnant les conditions initiales $x_0 = 1$ et $y_0 = 1$, on obtient l'hyperbole qui passe par le point (1 ; 1) du plan Oxy (fig. 222).

Remarquons qu'il ne passe aucune courbe intégrale par les points de l'axe Oy, autrement dit, ces points sont les points singuliers de cette équation.

4. Interprétation géométrique d'une équation. Soit donnée une équation différentielle du premier ordre $y' = f(x, y)$ et soit $y = \varphi(x)$ une solution. Le graphique de cette solution est une courbe intégrale continue admettant une tangente en chacun de ses points. De cette équation il s'ensuit que le coefficient directeur y' de la tangente à la courbe intégrale en chacun de ses points $(x; y)$ est égal à la valeur prise en ce point par le second membre $f(x, y)$ de l'équation. Donc l'équation $y' = f(x, y)$ établit une relation entre les coordonnées du point $(x; y)$ et le coefficient directeur y' de la tangente à la courbe intégrale en ce point. Sachant x et y, on peut définir la direction de la tangente à cette courbe en $(x; y)$.

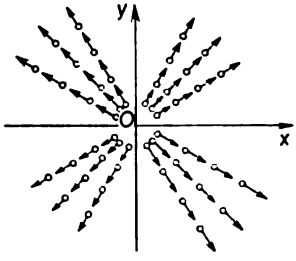


Fig. 223

Associons à chaque point $(x; y)$ de la courbe intégrale un segment orienté de coefficient directeur $f(x, y)$. On obtient le *champ de directions* de l'équation qui donne la signification géométrique de l'équation différentielle du premier ordre.

Ainsi l'équation $y' = f(x, y)$ définit sur le plan Oxy un champ de directions et toute solution de cette équation est une courbe intégrale admettant en chacun de ses points une tangente dont la direction est confondue avec celle du champ en ce point.

Le champ de directions d'une équation différentielle nous permet de construire approximativement les courbes intégrales de cette équation.

Exemple 3. Soit l'équation $y' = y/x$.

La fonction $f(x, y) = y/x$ n'étant pas définie pour $x = 0$, on peut construire le champ de directions sur le plan tout entier privé de l'axe Oy.

En chaque point $(x; y)$ $(x \neq 0)$ le coefficient directeur y' de la tangente à la courbe intégrale est égal à y/x et est confondu avec celui de la droite passant par ce point et par l'origine des coordonnées. Le champ de directions de cette équation est représenté sur la figure 223. Il est évident que les courbes intégrales sont les droites $y = Cx$, où C est une constante arbitraire.

Considérons maintenant des méthodes d'intégration des équations différentielles du premier ordre. Signalons qu'il n'existe pas de méthode générale d'intégration. Les équations différentielles sont réparties en types justiciables chacun d'une méthode de résolution.

5. Equations à variables séparables.

Définition 5. On appelle *équation différentielle à variables séparables* une équation de la forme

$$y' = f_1(x)\, f_2(y), \tag{6}$$

où $f_1(x)$ et $f_2(y)$ sont des fonctions continues.

Pour intégrer l'équation (6), il faut séparer les variables. A cet effet, remplaçons y' par $\frac{dy}{dx}$, divisons les deux membres de (6) par $f_2(y)$ (à condition que $f_2(y) \neq 0$) et multiplions par dx. L'équation (6) devient

$$\frac{dy}{f_2(y)} = f_1(x)\, dx. \tag{7}$$

Dans cette équation les variables sont séparées, c'est-à-dire qu'au premier membre figurent seulement des y et au second, seulement des x.

En admettant que $y = \varphi(x)$ est une solution et en la portant dans (7), on obtient une identité. L'intégration de cette identité nous donne

$$\int \frac{dy}{f_2(y)} + C_1 = \int f_1(x)\, dx + C_2, \text{ ou } \int \frac{dy}{f_2(y)} = \int f_1(x)\, dx + C, \text{ *)} \tag{8}$$

où $C = C_2 - C_1$ est une constante arbitraire.

La relation (8) définit implicitement la solution générale de l'équation (6).

Exemple 4. Intégrer l'équation $y' = y/x$ (comparer avec l'exemple 3).

Solution. Cette équation est de la forme (6), où $f_1(x) = 1/x$ et $f_2(y) = y$. En séparant les variables, on obtient $\frac{dy}{y} = \frac{dx}{x}$. Une intégration nous donne

$$\int \frac{dy}{y} = \int \frac{dx}{x} + \ln|C_1|, \; C_1 \neq 0 \text{ **)} \text{ ou } \ln|y| = \ln|x| + \ln|C_1|.$$

D'où $|y| = |C_1|\,|x|$, ce qui équivaut à l'équation $y = \pm C_1 x$.

En posant $\pm C_1 = C$, on obtient finalement la solution générale

$$y = Cx, \tag{9}$$

*) En théorie des équations différentielles le symbole de l'intégrale indéfinie ne désigne pas l'ensemble de toutes les primitives, mais l'une d'elles seulement.

**) Pour alléger l'écriture on a désigné la constante arbitraire non pas par C mais par $\ln|C_1|$, ce qui est possible, puisque $\ln|C_1|$ peut prendre toute valeur de l'intervalle $]-\infty, +\infty[$.

où C est une constante arbitraire non nulle. Remarquons que $y = 0$ est aussi une solution de l'équation (que nous avons perdue en divisant par y). On peut inclure cette solution dans (9) en faisant prendre la valeur 0 à la constante C. La solution générale (9) représente une famille de droites passant par l'origine des coordonnées.

Soit à déterminer la solution particulière qui satisfait aux conditions initiales $x_0 = 1$ et $y_0 = 2$. En substituant ces valeurs à x et y dans la solution générale (9), on obtient $2 = C \cdot 1$, d'où $C = 2$. La solution particulière cherchée est donc $y = 2x$.

6. Equations linéaires.

Définition 6. On appelle *équation différentielle linéaire du premier ordre* une équation de la forme

$$y' + p(x)\, y = f(x), \tag{10}$$

où $p(x)$ et $f(x)$ sont des fonctions continues.

Ces équations sont dites linéaires parce que la fonction inconnue y et sa dérivée y' y figurent linéairement, c'est-à-dire au premier degré.

Si $f(x) \equiv 0$, l'équation (10) s'appelle *équation linéaire sans second membre* ou *équation linéaire homogène*, et si $f(x) \not\equiv 0$, *équation linéaire avec second membre* ou *équation linéaire non homogène* ou encore *équation complète*.

La solution générale de l'équation (10) peut être déterminée par la méthode de variation de la constante.

Le principe de cette méthode est le suivant : on trouve d'abord la solution générale de l'équation linéaire sans second membre

$$y' + p(x)\, y = 0, \tag{11}$$

associée à l'équation complète (10). L'équation (11) est une équation à variables séparables. En séparant les variables et en intégrant on obtient

$$\frac{dy}{y} = -p(x)\,dx, \quad \ln|y| = -\int p(x)\,dx + \ln|C_1|,$$

d'où l'on déduit la solution générale de l'équation (11)

$$y = \pm C_1 e^{-\int p(x)\,dx}, \text{ ou } y = Ce^{-\int p(x)\,dx}, \tag{12}$$

où $C = \pm C_1$ est une constante arbitraire.

Cherchons maintenant la solution générale de (10) sous la forme (12), où C est traitée non plus comme une constante mais comme une fonction inconnue de x (d'où le nom de la méthode), c'est-à-dire sous la forme

$$y = C(x)\, e^{-\int p(x)\,dx}. \tag{13}$$

Pour déterminer $C(x)$ portons (13) dans (10). On obtient

$$C'(x)\,e^{-\int p(x)\,dx} - C(x)\,p(x)\,e^{-\int p(x)\,dx} + \\ + p(x)\,C(x)\,e^{-\int p(x)\,dx} = f(x)$$

ou

$$C'(x) = f(x)\,e^{\int p(x)\,dx}. \qquad (14)$$

Pour que la fonction (13) soit solution de l'équation (10), il faut que $C(x)$ soit solution de l'équation (14). Une intégration de (14) nous donne

$$C(x) = \int f(x)\,e^{\int p(x)\,dx}\,dx + C_1,$$

où C_1 est une constante arbitraire. En portant l'expression de $C(x)$ dans (13), on obtient la solution générale de l'équation linéaire (10) :

$$y(x) = C_1 e^{-\int p(x)\,dx} + e^{-\int p(x)\,dx} \int f(x)\,e^{\int p(x)\,dx}\,dx. \qquad (15)$$

La formule (15) est d'un usage compliqué et plutôt que de s'en servir directement il vaut mieux reprendre à chaque fois les raisonnements qui ont permis de l'établir.

Exemple 5. Trouver la solution générale de l'équation $y' + 3y = e^{2x}$.

Solution. Cette équation est linéaire. Ici $p(x) = 3$ et $f(x) = e^{2x}$. Intégrons tout d'abord l'équation sans second membre $y' + 3y = 0$. En séparant les variables $\frac{dy}{y} = -3\,dx$ et en intégrant, on trouve

$$\ln|y| = -3x + \ln|C_1| \text{ ou } y = \pm C_1 e^{-3x} = Ce^{-3x}.$$

Cherchons la solution générale de cette équation sous la forme $y = C(x)\,e^{-3x}$. Une dérivation nous donne $y' = C'(x)\,e^{-3x} - 3C(x)\,e^{-3x}$. En portant les expressions de y et y' dans l'équation initiale, on obtient

$$C'(x)\,e^{-3x} = e^{2x}, \quad C'(x) = e^{5x} \text{ ou } dC = e^{5x}\,dx,$$

d'où $C(x) = \frac{1}{5}e^{5x} + C_2$, où C_2 est une constante arbitraire. Donc la solution générale de cette équation est

$$y = C(x)\,e^{-3x} = \left(\frac{1}{5}e^{5x} + C_2\right)e^{-3x} \text{ ou } y = \frac{1}{5}e^{2x} + C_2 e^{-3x}.$$

7. Equations aux différentielles totales.

Définition 7. On appelle *équation aux différentielles totales* une équation de la forme

$$P(x, y)\,dx + Q(x, y)\,dy = 0, \tag{16}$$

où le premier membre est la différentielle totale d'une fonction $F(x, y)$ dans un domaine G.

Si l'équation (16) est une équation aux différentielles totales, on peut la mettre sous la forme

$$dF(x, y) = 0,$$

où $F(x, y)$ est une fonction telle que $dF(x, y) = P(x, y)\,dx + Q(x, y)\,dy$. D'où il s'ensuit que la solution générale de l'équation (16) est implicitement définie par l'équation

$$F(x, y) = C,$$

où C est une constante arbitraire. En effet, si $y = \varphi(x)$ est une solution de l'équation (16), alors $dF(x, \varphi(x)) \equiv 0$, c'est-à-dire que $F(x, \varphi(x)) = C$, et inversement, pour toute fonction $y = \varphi(x)$ transformant l'équation $F(x, y) = C$ en identité, on obtient $dF(x, \varphi(x)) = 0$, c'est-à-dire que $y = \varphi(x)$ est une solution de l'équation (16).

Donc l'intégration de l'équation (16) revient à chercher une fonction $F(x, y)$ dont la différentielle totale est égale à $P(x, y)\,dx + Q(x, y)\,dy$.

On sait (cf. théorème 13.7) qu'une condition nécessaire et suffisante pour que l'expression $P(x, y)\,dx + Q(x, y)\,dy$ soit la différentielle totale d'une fonction $F(x, y)$ est que

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}. \tag{17}$$

Supposons que la condition (17) est remplie. Il existe alors une fonction $F(x, y)$ telle que $dF = \frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial y}\,dy = P(x, y)\,dx + Q(x, y)\,dy$. D'où

$$\frac{\partial F}{\partial x} = P(x, y), \quad \frac{\partial F}{\partial y} = Q(x, y). \tag{18}$$

En intégrant l'expression $\frac{\partial F}{\partial x} = P(x, y)$ par rapport à x on trouve

$$F(x, y) = \int P(x, y)\,dx + C(y), \tag{19}$$

où $C(y)$ est une fonction arbitraire de y. Choisissons maintenant la fonction $C(y)$ de telle sorte que soit vérifiée la deuxième relation (18). Pour cela dérivons le second membre de (19) par rapport à y et égalons le résultat obtenu à $Q(x, y)$:

$$\frac{\partial}{\partial y}\left(\int P(x, y)\,dx\right) + C'(y) = Q(x, y). \tag{20}$$

Tirons $C'(y)$ de l'équation (20) et intégrons pour trouver $C(y)$. En portant la fonction $C(y)$ dans (19), on obtient la fonction cherchée $F(x, y)$.

Si l'on veut trouver la solution particulière qui satisfait aux conditions initiales $x = x_0$, $y = y_0$, il faut substituer x_0 et y_0 à x et y dans la solution générale $F(x, y) = C$. On a alors $C = F(x_0, y_0)$ et $F(x, y) = F(x_0, y_0)$ est la solution particulière cherchée.

Exemple 6. Trouver la solution générale de l'équation $(x + y + 1)\,dx + (x - y^2 + 3)\,dy = 0$ et la solution particulière qui satisfait aux conditions initiales $x_0 = 1$, $y_0 = 0$.

Solution. On a ici $P(x, y) = x + y + 1$, $Q(x, y) = x - y^2 + 3$. Puisque

$$\frac{\partial (x+y+1)}{\partial y} = 1 = \frac{\partial (x-y^2+3)}{\partial x},$$

l'expression $(x + y + 1)\,dx + (x - y^2 + 3)\,dy$ est la différentielle totale d'une fonction $F(x, y)$. On a

$$F(x, y) = \int P(x, y)\,dx + C(y) = \int (x+y+1)\,dx + C(y) = \frac{x^2}{2} + xy + x + C(y). \quad (21)$$

Trouvons la fonction $C(y)$ à l'aide de la formule (20) :

$$\frac{\partial}{\partial y}\left(\frac{x^2}{2} + xy + x + C(y)\right) = x - y^2 + 3;$$

$$C'(y) = -y^2 + 3; \quad \frac{dC}{dy} = -y^2 + 3; \quad dC = (-y^2 + 3)\,dy,$$

$$C(y) = \int (-y^2 + 3)\,dy + C_1 = -\frac{y^3}{3} + 3y + C_1.$$

En portant l'expression de $C(y)$ dans (21), on trouve

$$F(x, y) = \frac{x^2}{2} + xy + x - \frac{y^3}{3} + 3y + C_1.$$

L'équation prend alors la forme $dF(x, y) = 0$ et sa solution générale est donnée par l'équation

$$F(x, y) = C_2 \quad \text{ou} \quad \frac{x^2}{2} + xy + x - \frac{y^3}{3} + 3y + C_1 = C_2.$$

En posant $6(C_2 - C_1) = C_3$ (où C_3 est une constante arbitraire), on obtient finalement l'équation qui définit implicitement la solution générale de l'équation différentielle initiale

$$3x^2 + 6xy + 6x - 2y^3 + 18y = C_3.$$

Trouvons maintenant la solution particulière qui satisfait aux conditions initiales $x_0 = 1$ et $y_0 = 0$. En substituant ces

valeurs dans la solution générale on obtient $3 \cdot 1 + 6 \cdot 1 = C_3$, d'où $C_3 = 9$. Donc la solution particulière cherchée est

$$3x^2 + 6xy + 6x - 2y^3 + 18y = 9.$$

8. Intégration approchée des équations différentielles du premier ordre par la méthode d'Euler. Nous avons envisagé plusieurs méthodes d'intégration exacte des équations différentielles du premier ordre. Si aucune d'elles ne conduit au résultat escompté ou si elles donnent lieu à des calculs inextricables, on se tourne vers les méthodes de résolution approchée des équations. Voyons la plus simple d'entre elles: la *méthode d'Euler*.

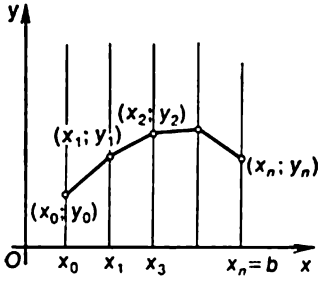


Fig. 224

L'idée de cette méthode est d'approcher la courbe intégrale cherchée qui est représentative d'une solution particulière par une ligne polygonale.

Soit donnée l'équation différentielle

$$y' = f(x, y)$$

avec les conditions initiales $y|_{x=x_0} = y_0$. Cherchons une solution approchée satisfaisant à ces conditions initiales sur un intervalle $[x_0, b]$.

Partageons l'intervalle $[x_0, b]$ en n parties égales par les points $x_0 < x_1 < x_2 < \ldots < x_n = b$. Posons $x_1 - x_0 = x_2 - x_1 = \ldots = x_n - x_{n-1} = \Delta x$. Désignons par y_i les valeurs approchées de la solution cherchée aux points x_i ($i = 1, 2, \ldots, n$). Menons par les points de subdivision x_i des droites parallèles à Oy (fig. 224) et effectuons successivement les opérations suivantes.

Portons les valeurs x_0 et y_0 dans le second membre de l'équation $y' = f(x, y)$ et calculons le coefficient directeur $y' = f(x_0, y_0)$ de la tangente à la courbe intégrale au point $(x_0; y_0)$. Pour trouver la valeur approchée y_1 de la solution cherchée remplaçons sur l'intervalle $[x_0, x_1]$ la courbe intégrale par un segment de sa tangente au point $(x_0; y_0)$. On obtient

$$y_1 - y_0 = f(x_0, y_0)(x_1 - x_0),$$

d'où, puisque x_0, x_1 et y_0 sont connus,

$$y_1 = y_0 + f(x_0, y_0)(x_1 - x_0) \quad \text{ou} \quad y_1 = y_0 + f(x_0, y_0)\,\Delta x.$$

En portant les valeurs x_1 et y_1 dans le second membre de l'équation $y' = f(x, y)$ on trouve le coefficient directeur $y' = f(x_1, y_1)$ de la tangente à la courbe intégrale au point $(x_1; y_1)$. En rempla-

çant sur l'intervalle $[x_1, x_2]$ la courbe intégrale par un segment de sa tangente, on obtient la valeur approchée y_2 au point x_2:

$$y_2 = y_1 + f(x_1, y_1)(x_2 - x_1) \quad \text{ou} \quad y_2 = y_1 + f(x_1, y_1)\,\Delta x.$$

Dans cette équation y_2 est exprimé en fonction de x_1, y_1 et x_2 qui sont connus.

En poursuivant cette procédure on trouve

$$y_3 = y_2 + f(x_2, y_2)\,\Delta x, \ \ldots, \ y_n = y_{n-1} + f(x_{n-1}, y_{n-1})\,\Delta x.$$

Nous avons ainsi construit approximativement la courbe intégrale cherchée sous forme d'une ligne polygonale et avons trouvé les valeurs approchées y_i aux points x_i à l'aide de la formule

$$y_i = y_{i-1} + f(x_{i-1}, y_{i-1})\,\Delta x \quad (i = 1, 2, \ldots, n). \qquad (22)$$

La formule (22) est la principale formule de calcul de la méthode d'Euler. Sa précision est d'autant plus élevée que la différence Δx est petite.

La méthode d'Euler n'est pas d'une précision très élevée. Il existe de bien meilleures méthodes d'intégration approchée des équations différentielles.

Exemple 7. Trouver une solution approchée de l'équation $y' = y + x$ sur l'intervalle [0, 1] satisfaisant aux conditions initiales $x_0 = 0$, $y_0 = 1$ et calculer y pour $x = 1$.

Solution. Partageons l'intervalle [0, 1] en 10 parties égales à l'aide des points $x_0 = 0$, $x_1 = 0{,}1$, $x_2 = 0{,}2$, ..., $x_{10} = 1$. Désignons par $y_1, y_2, \ldots, y_{10}$ les valeurs approchées de la solution que nous cherchons à l'aide de la formule (22). On a

$$y_1 = 1 + (1 + 0)\cdot 0{,}1 = 1{,}1; \quad y_2 = 1{,}1 + (1{,}1 + 0{,}1)\cdot 0{,}1 = 1{,}22.$$

On trouve de façon analogue les autres valeurs y_i. Consignons les résultats dans le tableau suivant:

x	y	$f(x, y)\,\Delta x$	x	y	$f(x, y)\,\Delta x$
$x_0 = 0$	$y_0 = 1$	0,1	$x_6 = 0{,}6$	$y_6 = 1{,}9431$	0,2543
$x_1 = 0{,}1$	$y_1 = 1{,}1$	0,12	$x_7 = 0{,}7$	$y_7 = 2{,}1974$	0,2897
$x_2 = 0{,}2$	$y_2 = 1{,}22$	0,142	$x_8 = 0{,}8$	$y_8 = 2{,}4871$	0,3287
$x_3 = 0{,}3$	$y_3 = 1{,}36$	0,1662	$x_9 = 0{,}9$	$y_9 = 2{,}8158$	0,3715
$x_4 = 0{,}4$	$y_4 = 1{,}5282$	0,1928	$x_{10} = 1$	$y_{10} = 3.1873$	
$x_5 = 0{,}5$	$y_5 = 1{,}7210$	0,2221			

La deuxième colonne contient les valeurs approchées y_i satisfaisant aux conditions initiales données sur [0, 1].

La valeur approchée de la solution pour $x = 1$ est $y_{10} = 3{,}1873$.

Pour avoir une idée de la précision de ces calculs intégrons cette équation pour les mêmes conditions initiales. Utilisons la méthode de variation de la constante puisque cette équation est linéaire. Cherchons la solution générale de l'équation sans second membre associée

$$\frac{dy}{dx} = y, \quad \frac{dy}{y} = dx, \quad \ln|y| = y + \ln|C|, \quad y = Ce^x.$$

Faisons varier la constante: $y = C(x)\,e^x$, $y' = C'(x)\,e^x + C(x)\,e^x$. En portant dans l'équation, on obtient: $C'(x)\,e^x + C(x)\,e^x = C(x)\,e^x + x$, $C'(x) = xe^{-x}$, $C(x) = -xe^{-x} - e^{-x} + C_1$. Donc

$$y = C_1 e^x - x - 1$$

est la solution générale. En portant $x_0 = 0$, $y_0 = 1$ dans cette solution, on trouve $C_1 = 2$. Ainsi la solution exacte qui satisfait aux conditions initiales données est $y = 2e^x - x - 1$. Pour $x = 1$, $y(1) = 2(e - 1) \approx 3{,}4366$. En comparant avec la valeur approchée on voit que l'erreur absolue est 0,2293.

9. Quelques applications des équations différentielles du premier ordre. De nombreux problèmes de physique conduisent aux équations différentielles du premier ordre. La principale difficulté est la mise en équations de ces problèmes, difficulté accrue par l'absence de toute méthode universelle. Chaque problème nécessite une approche individuelle impliquant une profonde compréhension des lois de la physique et une maîtrise de l'art de transcrire les problèmes de physique dans le langage des mathématiques.

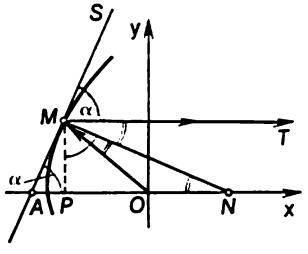


Fig. 225

Voyons quelques problèmes.

Problème du projecteur. Construire un miroir ayant la forme d'une surface de révolution reflétant les rayons issus d'une source placée en un point O de l'axe de rotation dans une direction parallèle à cet axe.

Pour résoudre ce problème considérons une section de ce miroir par un plan passant par l'axe de rotation. Plaçons la source de lumière en l'origine des coordonnées et supposons que l'axe de rotation est l'axe Ox (fig. 225). Désignons par α l'angle de l'axe Ox et de la tangente AS en un point arbitraire $M(x; y)$ de la section. On se propose de déterminer la forme de cette section, c'est-à-dire d'établir une relation entre y et x. Le chemin OMT est le chemin d'un rayon issu de O et reflété parallèlement à l'axe

Ox. Traçons la normale MN et abaissons la perpendiculaire MP sur Ox. Puisque $\widehat{AMO} = \widehat{TMS}$ (l'angle d'incidence est égal à l'angle de réflexion), on a $\widehat{OMN} = \widehat{NMT} = \widehat{ONM}$. Donc le triangle NOM est isocèle et $OM = ON$. D'autre part, $\widehat{PMN} = \alpha$ par construction. En se servant des égalités $ON = PN - PO$, $PN = y \operatorname{tg} \alpha$, $PO = -x$, $ON = OM = \sqrt{x^2 + y^2}$, on obtient la relation

$$ON = y \operatorname{tg} \alpha + x = \sqrt{x^2 + y^2}.$$

Puisque $\frac{dy}{dx} = \operatorname{tg} \alpha$, la relation que l'on cherche à établir entre y et x est l'équation différentielle du premier ordre

$$x + \frac{dy}{dx}\, y = \sqrt{x^2 + y^2}.$$

Transformons cette équation de la manière suivante. Multiplions ses deux membres par $2dx$:

$$2x\,dx + 2y\,dy = 2\sqrt{x^2 + y^2}\,dx \quad \text{ou} \quad d(x^2 + y^2) = 2\sqrt{x^2 + y^2}\,dx.$$

La substitution $z = x^2 + y^2$ nous conduit à l'équation à variables séparables

$$dz = 2\sqrt{z}\,dx$$

ou

$$z^{-1/2}\,dz = 2dx, \text{ d'où } \sqrt{z} = x + C.$$

Comme $z = x^2 + y^2$, il vient

$$\sqrt{x^2 + y^2} = x + C.$$

En élevant les deux membres au carré, on trouve

$$y^2 = 2C\,(x + C/2).$$

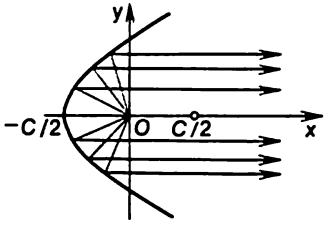


Fig. 226

Donc la courbe cherchée est une parabole de paramètre $p = C$ dont le sommet est situé sur l'axe Ox et a pour abscisse $-C/2$ (fig. 226). Donc la surface de révolution cherchée est un *paraboloïde de révolution.*

Désintégration radio-active. On a établi empiriquement que la vitesse de désintégration d'une matière radio-active, c'est-à-dire la vitesse de variation de sa masse en fonction du temps, était directement proportionnelle à sa quantité. Etablissons la loi de variation d'une masse m de matière radio-active en fonction du temps t sachant que $m = m_0$ pour $t = 0$.

Supposons que la masse de matière est m à l'instant t et $m - \Delta m$ à l'instant $t + \Delta t$. La quantité Δm se désintègre en un

temps Δt. Le rapport $\frac{\Delta m}{\Delta t}$ est la vitesse moyenne de désintégration durant Δt, et $\lim\limits_{\Delta t \to 0} \frac{\Delta m}{\Delta t} = \frac{dm}{dt}$, la vitesse instantanée de désintégration à l'instant t. Par hypothèse,

$$\frac{dm}{dt} = -km, \tag{23}$$

où k est le coefficient de proportionnalité (le signe « — » indique que la masse diminue avec le temps et la dérivée d'une fonction strictement décroissante est strictement négative). On a obtenu une équation différentielle du premier ordre dont l'intégration nous donnera une relation entre m et t.

On a

$$\frac{dm}{m} = -k\,dt, \quad \ln m = -kt + \ln C,$$

d'où

$$m = Ce^{-kt}. \tag{24}$$

La formule (24) exprime la masse de matière en fonction du temps. Dans ce problème la constante C a une valeur bien définie, plus exactement, pour $t = 0$, on obtient $m_0 = Ce^0 = C$. En portant cette valeur de C dans la formule (24), on obtient la loi de variation de la masse de matière radio-active en fonction du temps :

$$m = m_0 e^{-kt}. \tag{25}$$

La formule (24) nous donne la solution générale de l'équation différentielle de la désintégration radio-active, la formule (25), la solution particulière satisfaisant aux conditions initiales données.

Le coefficient k se détermine expérimentalement. Pour le radium, par exemple, $k \approx 0{,}000447$. L'intervalle de temps durant lequel la quantité initiale de matière radio-active diminue de moitié s'appelle *période de désintégration* de cette matière. En remplaçant dans (25) m par $m_0/2$ et k par sa valeur 0,000447, on obtient une équation pour la détermination de la période de désintégration T du radium :

$$\frac{m_0}{2} = m_0 e^{-0{,}000447T}, \quad -0{,}000447T = -\ln 2,$$

d'où

$$T = \frac{\ln 2}{0{,}000447}; \quad T \approx 1550 \text{ années}.$$

Loi de croissance normale. La loi de croissance normale est une loi en vertu de laquelle la vitesse de croissance de la matière est directement proportionnelle à sa quantité. Trouvons la for-

mule qui donne la variation de la quantité de matière y en fonction du temps t sachant que $y = y_0$ à l'instant $t = 0$.

La variable indépendante est ici le temps t et l'inconnue, la quantité de matière à tout instant. La vitesse de croissance est la vitesse de variation de y en fonction de t.

En utilisant comme plus haut la signification physique de la dérivée on peut mettre la loi de croissance normale sous la forme

$$\frac{dy}{dt} = ky, \tag{26}$$

où $k > 0$ est le coefficient de proportionnalité. L'équation (26) diffère seulement par son signe de l'équation (23) qui décrit de nombreux processus de reproduction.

La solution de l'équation (26) qui satisfait aux conditions initiales $t = 0$, $y = y_0$ est

$$y = y_0 e^{kt}. \tag{27}$$

La formule (27) traduit précisément la loi de croissance normale. Cette loi régit par exemple la multiplication des neutrons dans un réacteur nucléaire, la reproduction des bactéries, la croissance des cristaux, etc.

§ 2. Equations différentielles du second ordre

1. Notions fondamentales.

Définition. On appelle *équation différentielle du second ordre* une équation de la forme

$$F(x, y, y', y'') = 0,$$

où x est la variable indépendante; y la fonction inconnue; y' et y'', ses dérivées.

On étudie généralement les équations qui peuvent être ramenées à la forme résolue par rapport à la dérivée seconde:

$$y'' = f(x, y, y'). \tag{1}$$

Comme dans le cas d'une équation différentielle du premier ordre, on appelle solution de l'équation (1) une fonction $y = \varphi(x)$, $x \in]a, b[$, qui transforme (1) en identité. Le graphique d'une solution s'appelle aussi *courbe intégrale*.

Les équations du second ordre sont justiciables aussi d'un théorème d'existence et d'unicité de la solution (théorème de Cauchy) identique au théorème correspondant pour les équations du premier ordre.

Théorème 15.2 (Cauchy). *Si une fonction $f(x, y, y')$ et ses dérivées partielles $f'_y(x, y, y')$ et $f'_{y'}(x, y, y')$ sont définies et con-*

tinues dans un domaine G de l'espace $(x; y; y')$, *alors quel que soit le point intérieur* $(x_0; y_0; y_0')$ *du domaine G, l'équation différentielle* $y'' = f(x, y, y')$ *admet dans un voisinage du point* x_0 *une seule solution satisfaisant aux conditions*

$$y = y_0, \; y' = y_0' \quad pour \quad x = x_0. \tag{2}$$

Géométriquement cela signifie que par le point $(x_0; y_0)$ du plan il passe une seule courbe intégrale dont la tangente en ce point a un coefficient directeur y_0' donné.

Les conditions (2) s'appellent conditions initiales et sont souvent notées sous la forme

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0'. \tag{3}$$

Comme pour l'équation du premier ordre, le problème de Cauchy est le problème qui consiste à trouver une solution vérifiant des conditions initiales données.

Définissons maintenant la solution générale et la solution particulière qui vérifie des conditions initiales données.

On dit qu'une fonction $y = \varphi(x, C_1, C_2)$ dépendant de x et de deux constantes arbitraires C_1 et C_2 est la *solution générale* de l'équation (1) dans un domaine G si elle est solution de l'équation (1) pour toutes valeurs des constantes C_1 et C_2 et si pour toutes conditions initiales (3) il existe des valeurs uniques C_1^0 et C_2^0 de C_1 et C_2 pour lesquelles la fonction $y = \varphi(x, C_1^0, C_2^0)$ vérifie les conditions initiales (3).

Toute fonction $y = \varphi(x, C_1^0, C_2^0)$ déduite de la solution générale $y = \varphi(x, C_1, C_2)$ de l'équation (1) en donnant des valeurs particulières C_1^0 et C_2^0 aux constantes C_1 et C_2 s'appelle *solution particulière.*

Soit par exemple l'équation $y'' = 2$. Cette équation est du second ordre. Les fonctions $f(x, y, y') = 2$, $f_y'(x, y, y') = 0$ et $f_{y'}'(x, y, y') = 0$ étant continues et définies sur l'espace $(x; y; y')$ tout entier, cette équation satisfait aux conditions du théorème de Cauchy sur l'espace tout entier.

La solution générale de cette équation s'obtient par une double intégration. La première intégration nous donne la dérivée première $y' = 2x + C_1$, la deuxième intégration, la solution générale

$$y = x^2 + C_1 x + C_2,$$

où C_1 et C_2 sont des constantes arbitraires. La solution générale représente une famille de paraboles et comme cette solution dépend de deux constantes arbitraires, par chaque point du plan il passe une infinité de paraboles présentant des tangentes distinctes en ce point. Donc pour particulariser une parabole de cette famille il faut en plus du point $(x_0; y_0)$ par lequel passe cette parabole se donner le coefficient directeur y_0' de la tangente à la parabole en ce point.

Trouvons par exemple la solution particulière de cette équation qui vérifie les conditions initiales $y|_{x=1} = 1$, $y'|_{x=1} = 1$. En portant ces valeurs dans l'expression de la solution générale $y = x^2 + C_1x + C_2$ et dans sa dérivée $y' = 2x + C_1$, on obtient le système d'équations suivant en C_1 et C_2:

$$\begin{cases} 1 = 1 + C_1 + C_2, \\ 1 = 2 + C_1, \end{cases}$$

d'où $C_1 = -1$ et $C_2 = 1$. Donc la solution particulière cherchée est

$$y = x^2 - x + 1$$

et son graphique est une parabole passant par le point (1; 1) et admettant en ce point une tangente de coefficient directeur 1.

2. Abaissement de l'ordre des équations du second ordre. Considérons trois cas particuliers où l'intégration de l'équation (1) se ramène à celle d'une équation du premier ordre par un changement de variables. Cette transformation de l'équation (1) s'appelle *abaissement de l'ordre*.

1) **Equation de la forme $y'' = f(x)$.** L'équation ne contient ni y, ni y'. Posons $z(x) = y'$. Alors $z'(x) = y''$ et l'équation donnée se transforme en une équation du premier ordre $z'(x) = f(x)$. Son intégration nous donne $z(x) = \int f(x)\,dx + C_1$. Comme $z(x) = y'$, il vient $y' = \int f(x)\,dx + C_1$. D'où

$$y = \int \left[\int f(x)\,dx\right] dx + C_1x + C_2,$$

où C_1 et C_2 sont des constantes arbitraires.

Exemple 1. Trouver la solution générale de l'équation $y'' = x$.

Solution. En posant $z(x) = y'$, on obtient l'équation du premier ordre $z'(x) = x$. Une intégration nous donne $z(x) = \frac{x^2}{2} + C_1$. En remplaçant $z(x)$ par y' et en intégrant on obtient la solution générale cherchée:

$$y = \int \left[\frac{x^2}{2} + C_1\right] dx + C_2 = \frac{x^3}{6} + C_1x + C_2.$$

2) **Equation de la forme $y'' = f(x, y')$.** L'équation ne contient pas y. Posons comme précédemment $z(x) = y'$; alors $z'(x) = y''$ et l'équation se transforme en l'équation du premier ordre $z' = f(x, z)$. En intégrant on trouve $z(x) = \varphi(x, C_1)$.

Comme $z(x) = y'$, il vient $y' = \varphi(x, C_1)$. D'où

$$y = \int \varphi(x, C_1)\,dx + C_2,$$

où C_1 et C_2 sont des constantes arbitraires.

Exemple 2. Trouver la solution générale de l'équation $y'' - \frac{3}{x}y' = x$.

Solution. En posant $z(x) = y'$, on obtient l'équation linéaire du premier ordre $z' - \frac{3}{x}z = x$. Son intégration nous donne $z(x) = C_1x^3 - x^2$. Alors $y' = C_1x^3 - x^2$ et la solution générale est $y = C_1\frac{x^4}{4} - \frac{x^3}{3} + C_2$.

3) **Equation de la forme $y'' = f(y, y')$.** L'équation ne contient pas x. Posons $y' = z(y)$. Alors

$$y'' = \frac{d(y')}{dx} = \frac{dy'}{dy}\frac{dy}{dx} = \frac{dz}{dy}\frac{dy}{dx} = \frac{dz}{dy}z(y).$$

En portant les expressions de y' et y'' dans l'équation, on obtient une équation du premier ordre en $z(y)$:

$$z\frac{dz}{dy} = f(y, z).$$

Son intégration nous donne $z = \varphi(y, C_1)$. Comme $z = \frac{dy}{dx}$, il vient $\frac{dy}{dx} = \varphi(y, C_1)$. D'où $\frac{dy}{\varphi(y, C_1)} = dx$. On a obtenu une équation à variables séparées dont la solution générale est

$$\int \frac{dy}{\varphi(y, C_1)} = x + C_2,$$

où C_1 et C_2 sont des constantes arbitraires.

Exemple 3. Trouver la solution générale de l'équation $yy'' - 2y'^2 = 0$.

Solution. Posons $y' = z(y)$. Alors $y'' = z\frac{dz}{dy}$ et $zy\frac{dz}{dy} - 2z^2 = 0$. On reconnaît ici une équation du premier ordre à variables séparables. En la ramenant à la forme $\frac{dz}{z} = \frac{2dy}{y}$ et en intégrant on obtient $\ln|z| = 2\ln|y| + \ln|C_1|$, d'où $z = C_1y^2$. Comme $z = \frac{dy}{dx}$, on trouve $\frac{dy}{y^2} = C_1\,dx$ et la solution générale est

$$-\frac{1}{y} = C_1x + C_2 \quad \text{ou} \quad y = -\frac{1}{C_1x + C_2}.$$

En simplifiant par z on a perdu la solution $z = y' = 0$, c'est-à-dire la solution $y = C = \text{const}$. Dans le cas envisagé cette solu-

tion figure dans la solution générale, puisqu'on l'obtient pour $C_1 = 0$ (la solution $y = 0$ est exclue).

3. Equations différentielles d'ordre supérieur. Bornons-nous aux définitions fondamentales et aux remarques générales.

Une équation différentielle d'ordre n est de la forme

$$F(x, y, y', \ldots, y^{(n)}) = 0$$

ou, si elle est résolue par rapport à la dérivée supérieure,

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)}). \tag{4}$$

Comme pour les équations du premier et du second ordre, on appelle solution de l'équation (4) une fonction $y = \varphi(x)$, $x \in]a, b[$ qui transforme cette équation en identité.

Le théorème d'existence et d'unicité de la solution de l'équation (4) est identique aux théorèmes correspondants énoncés pour $n = 1$ et $n = 2$.

La solution générale de l'équation (4) dépend de x et de n constantes arbitraires et peut être mise sous la forme

$$y = \varphi(x, C_1, C_2, \ldots, C_n).$$

Les solutions qui se déduisent de la solution générale pour des valeurs particulières des constantes $C_1, C_2, \ldots, C_n$ s'appellent solutions particulières de l'équation (4). Pour obtenir une solution particulière à partir de la solution générale, il faut se donner les conditions initiales

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0', \quad y''|_{x=x_0} = y_0'', \ldots, y^{(n-1)}|_{x=x_0} = y_0^{(n-1)}. \tag{5}$$

La recherche de la solution de l'équation (4) qui vérifie les conditions initiales (5) s'appelle problème de Cauchy pour cette équation.

Un exemple élémentaire d'équation de la forme (4) est fourni par une équation dont le second membre dépend uniquement de x, c'est-à-dire par une équation de la forme

$$y^{(n)} = f(x). \tag{6}$$

Cette équation s'intègre sans peine. En effet, en intégrant n fois de suite, on obtient

$$y^{(n-1)} = \int f(x\ \mathrm{d}x + C_1,$$

$$y^{(n-2)} = \int \left[\int f(x)\,\mathrm{d}x + C_1 \right] \mathrm{d}x + C_2 = \int \mathrm{d}x \int f(x)\,\mathrm{d}x + C_1 x + C_2,$$

. . . . . .

$$y = \int \mathrm{d}x \int \mathrm{d}x \ldots \int f(x)\,\mathrm{d}x + C_1 \frac{x^{n-1}}{(n-1)!} + C_2 \frac{x^{n-2}}{(n-2)!} + \ldots + C_n, \tag{7}$$

où $C_1, C_2, \ldots, C_n$ sont des constantes arbitraires. La fonction (7) est précisément la solution générale de l'équation (6).

Exemple 4. Trouver la solution générale de l'équation du troisième ordre $y''' = e^x$ et la solution particulière qui vérifie les conditions initiales: $y|_{x=0} = 0$, $y'|_{x=0} = 0$, $y''|_{x=0} = 1$.

Solution. Une double intégration nous donne: $y'' = e^x + C_1$, $y' = e^x + C_1 x + C_2$. En intégrant encore on obtient la solution générale $y = e^x + C_1 \frac{x^2}{2} + C_2 x + C_3$, où C_1, C_2 et C_3 sont des constantes arbitraires.

En portant les conditions initiales dans les expressions de y, y' et y'', on obtient $0 = 1 + C_3$, $0 = 1 + C_2$, $1 = 1 + C_1$, d'où $C_1 = 0$, $C_2 = C_3 = -1$. Donc la solution particulière demandée est $y = e^x - x - 1$.

§ 3. Equations différentielles linéaires du second ordre

En théorie des équations différentielles, les équations différentielles linéaires du second ordre occupent une place privilégiée non seulement parce qu'elles constituent un type d'équations simple et bien étudié, mais encore parce qu'elles décrivent de nombreuses lois en physique, mécanique et particulièrement en électrotechnique.

1. Notions fondamentales.

Définition. On appelle *équation différentielle linéaire du second ordre* une équation de la forme

$$y'' + p(x)\, y' + q(x)\, y = f(x), \tag{1}$$

où y est la fonction inconnue, $p(x)$, $q(x)$ et $f(x)$ des fonctions continues sur un intervalle $]a, b[$.

Si $f(x) \equiv 0$, l'équation (1) s'appelle *équation linéaire homogène*, ou *sans second membre*; si $f(x) \not\equiv 0$, *équation linéaire avec second membre*, ou *non homogène*, ou *complète*.

En résolvant l'équation (1) par rapport à y'', on obtient $y'' = -p(x)\, y' - q(x)\, y + f(x)$. On voit que cette équation est un cas particulier de l'équation $y'' = f(x, y, y')$ et qu'elle vérifie les conditions du théorème d'existence et d'unicité de la solution. En effet, la fonction $f(x, y, y') = -p(x)\, y' - q(x)\, y + f(x)$ est continue par rapport aux variables x, y et y' (elle dépend linéairement de y et y' et les fonctions $p(x)$, $q(x)$ et $f(x)$ sont continues par hypothèse) et ses dérivées partielles $f'_y(x, y, y') = -q(x)$ et $f'_{y'}(x, y, y') = -p(x)$ aussi ($p(x)$ et $q(x)$ ne dépendent pas de y et y' et sont continues par rapport à x par hypothèse). Donc, quelles que soient les conditions initiales $y|_{x=x_0} = y_0$, $y'|_{x=x_0} =$

$= y_0'$, où $x_0 \in]a, b[$, l'équation (1) admet une seule solution qui les vérifie.

Commençons l'étude des équations différentielles linéaires par les équations homogènes.

2. Equations différentielles linéaires sans second membre du second ordre. Considérons certaines propriétés des solutions des équations différentielles linéaires sans second membre.

Théorème 15.3. *Si des fonctions $y_1(x)$ et $y_2(x)$ sont solutions de l'équation*

$$y'' + p(x)\, y' + q(x)\, y = 0, \tag{2}$$

il en est de même de la fonction $y = C_1 y_1(x) + C_2 y_2(x)$ quelles que soient les constantes C_1 et C_2.

Démonstration. En dérivant deux fois la fonction $y = C_1 y_1(x) + C_2 y_2(x)$ et en portant les expressions de y, y' et y'' dans le premier membre de l'équation (2), on obtient

$$C_1 y_1''(x) + C_2 y_2''(x) + p(x)\,(C_1 y_1'(x) + C_2 y_2'(x)) +$$
$$+ q(x)\,(C_1 y_1(x) + C_2 y_2(x)) = C_1 [y_1''(x) + p(x)\, y_1'(x) +$$
$$+ q(x)\, y_1(x)] + C_2 [y_2''(x) + p(x)\, y_2'(x) + q(x)\, y_2(x)].$$

Les fonctions $y_1(x)$ et $y_2(x)$ étant par hypothèse solutions de l'équation (2), les crochets sont identiquement nuls, ce qui exprime que la fonction $y = C_1 y_1(x) + C_2 y_2(x)$ est solution de l'équation (2). ■

Ainsi toute fonction de la forme $y = C_1 y_1(x) + C_2 y_2(x)$, où C_1 et C_2 sont des constantes arbitraires, est solution de l'équation (2). On peut se demander naturellement si cette solution n'est pas la solution générale de l'équation (2). Il s'avère que c'est le cas mais sous certaines conditions. Pour les établir introduisons la notion de dépendance linéaire et d'indépendance linéaire des fonctions $y_1(x)$ et $y_2(x)$.

On dit que des fonctions $y_1(x)$ et $y_2(x)$ sont *linéairement dépendantes* sur $]a, b[$ s'il existe des constantes α_1 et α_2 non toutes deux nulles telles que pour tout $x \in]a, b[$

$$\alpha_1 y_1(x) + \alpha_2 y_2(x) = 0. \tag{3}$$

Il est évident que si les fonctions $y_1(x)$ et $y_2(x)$ sont linéairement dépendantes, elles sont proportionnelles. En effet, si $\alpha_1 y_1(x) + \alpha_2 y_2(x) = 0$ et de plus $\alpha_1 \neq 0$ et $y_2(x) \neq 0$, alors $\frac{y_1(x)}{y_2(x)} = -\frac{\alpha_2}{\alpha_1} = \text{const}$. La réciproque est vraie.

On dit que des fonctions $y_1(x)$ et $y_2(x)$ sont *linéairement indépendantes* sur $]a, b[$ si la relation (3) entraîne $\alpha_1 = \alpha_2 = 0$.

Il est évident que si des fonctions $y_1(x)$ et $y_2(x)$ sont linéairement indépendantes, leur rapport $\frac{y_1(x)}{y_2(x)} \neq \text{const}$, c'est-à-dire qu'elles ne sont pas proportionnelles. Ainsi, par exemple, les fonctions $y_1(x) = x^2$ et $y_2(x) = x^3$ sont linéairement indépendantes sur tout intervalle $]a, b[$, puisque $\frac{y_1(x)}{y_2(x)} = \frac{1}{x} \neq \text{const}$, et les fonctions $y_1(x) = 4x^2$ et $y_2(x) = x^2$ sont linéairement dépendantes sur tout intervalle, puisque $\frac{y_1(x)}{y_2(x)} = 4 = \text{const}$.

Supposons maintenant que des fonctions $y_1(x)$ et $y_2(x)$ sont solutions de l'équation (2). La réponse à la question de savoir si elles sont linéairement dépendantes ou indépendantes nous est fournie par le *déterminant de Wronski* *), ou *wronskien*

$$\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1 y_2' - y_2 y_1'. \tag{4}$$

Le wronskien est une fonction définie sur $]a, b[$ qui est désignée par $W(y_1, y_2)$ ou simplement par $W(x)$.

Théorème 15.4. *Si des fonctions $y_1(x)$ et $y_2(x)$ sont linéairement dépendantes sur un intervalle $]a, b[$, leur wronskien est nul sur cet intervalle.*

Démonstration. Les fonctions $y_1(x)$ et $y_2(x)$ étant linéairement dépendantes, il existe par définition des nombres α_1 et α_2 non tous deux nuls tels que soit réalisée (3). Supposons pour fixer les idées que $\alpha_1 \neq 0$. De l'égalité (3) on déduit

$$y_1(x) = -\frac{\alpha_2}{\alpha_1} y_2(x), \quad y_1'(x) = -\frac{\alpha_2}{\alpha_1} y_2'(x).$$

En portant ces expressions de $y_1(x)$ et $y_1'(x)$ dans le wronskien on obtient

$$W(x) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = \begin{vmatrix} -\frac{\alpha_2}{\alpha_1} y_2 & y_2 \\ -\frac{\alpha_2}{\alpha_1} y_2' & y_2' \end{vmatrix} = 0. \quad \blacksquare$$

Théorème 15.5. *Si des solutions $y_1(x)$ et $y_2(x)$ de l'équation (2) sont linéairement indépendantes sur $]a, b[$, leur wronskien est non nul sur cet intervalle.*

Démonstration. Supposons par absurde qu'il existe un point $x_0 \in]a, b[$ en lequel $W(x_0) = 0$. Formons le système

*) Wronski Joseph (1775-1853), mathématicien polonais.

d'équations

$$\begin{cases} \alpha_1 y_1(x_0) + \alpha_2 y_2(x_0) = 0, \\ \alpha_1 y_1'(x_0) + \alpha_2 y_2'(x_0) = 0, \end{cases} \tag{5}$$

dans lequel α_1 et α_2 sont des inconnues. Le déterminant $W(x_0)$ de ce système étant nul, celui-ci admet (chap. X, § 3, cas 2) une solution non nulle (c'est-à-dire que l'un au moins des nombres α_1 et α_2 est non nul). Considérons la fonction

$$y(x) = \alpha_1 y_1(x) + \alpha_2 y_2(x),$$

où α_1, α_2 est une solution non triviale du système (5).

D'après le théorème 15.3, cette fonction est solution de l'équation (2). De plus, puisque α_1, α_2 est une solution du système (5), la fonction $y(x)$ satisfait aux conditions initiales nulles

$$y|_{x=x_0} = 0, \quad y'|_{x=x_0} = 0. \tag{6}$$

Ces conditions initiales sont de toute évidence vérifiées par la solution $y(x) = 0$. D'après le théorème d'existence et d'unicité, la solution $y(x) = 0$ est l'unique solution de l'équation (2) satisfaisant aux conditions (6). Donc $\alpha_1 y_1(x) + \alpha_2 y_2(x) = 0$ sur l'intervalle $]a, b[$, ce qui exprime que les fonctions $y_1(x)$ et $y_2(x)$ sont linéairement dépendantes sur $]a, b[$. Or ceci contredit le théorème. Donc $W(x) \neq 0$ pour tout $x \in]a, b[$. ■

On a ainsi établi que si des fonctions $y_1(x)$ et $y_2(x)$ sont solutions d'une équation linéaire sans second membre (2) sur $]a, b[$, alors leur wronskien est nul sur $]a, b[$ si elles sont linéairement dépendantes, et non nul si elles sont linéairement indépendantes.

Voyons maintenant sous quelles conditions la fonction $y = C_1 y_1(x) + C_2 y_2(x)$ est solution générale de l'équation (2).

Théorème 15.6. *Si des fonctions $y_1(x)$ et $y_2(x)$ linéairement indépendantes sur $]a, b[$ sont solutions de l'équation* (2), *la fonction*

$$y = C_1 y_1(x) + C_2 y_2(x), \tag{7}$$

où C_1 et C_2 sont des constantes arbitraires, est la solution générale de l'équation (2).

Démonstration. Rappelons qu'en vertu du théorème 15.3 la fonction $y = C_1 y_1(x) + C_2 y_2(x)$ est solution de l'équation (2) quelles que soient les constantes C_1 et C_2. Pour prouver qu'elle est solution générale, il suffit de montrer qu'on peut en déduire toute solution particulière satisfaisant à des conditions initiales quelconques.

Soit $x_0 \in]a, b[$ et soient

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0' \tag{8}$$

des conditions initiales arbitraires. Montrons que les constantes C_1 et C_2 peuvent être choisies de telle sorte que la solution (7) correspondante soit la solution particulière qui vérifie les conditions initiales (8).

Formons le système d'équations

$$\begin{cases} y_0 = C_1 y_1(x_0) + C_2 y_2(x_0), \\ y_0' = C_1 y_1'(x_0) + C_2 y_2'(x_0), \end{cases} \tag{9}$$

où C_1 et C_2 sont des inconnues. Le déterminant de ce système est le wronskien $W(x_0)$. Ce wronskien est $\neq 0$ puisque les fonctions $y_1(x)$ et $y_2(x)$ sont linéairement indépendantes sur $]a, b[$. Donc le système (9) admet une solution unique que nous désignerons par $C_1 = C_1^0$, $C_2 = C_2^0$. En portant C_1^0 et C_2^0 dans (7), on obtient la solution particulière de l'équation (2) qui satisfait aux conditions initiales (8), à savoir $y = C_1^0 y_1(x) + C_2^0 y_2(x)$. Ce qui exprime que la solution (7) est la solution générale de l'équation (2). ■

D'après le théorème prouvé, pour trouver la solution générale de l'équation (2), il suffit de connaître deux solutions particulières linéairement indépendantes et de former l'expression (7) avec des constantes arbitraires C_1 et C_2.

Exemple 1. Trouver la solution générale de l'équation $y'' - y = 0$.

Solution. Il est immédiat de remarquer que $y_1(x) = e^x$ et $y_2(x) = e^{-x}$ sont des solutions particulières. Le wronskien

$$W(x) = \begin{vmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{vmatrix} = -2$$

étant différent de 0, ces solutions sont linéairement indépendantes sur la droite numérique tout entière. Donc la solution générale est $y = C_1 e^x + C_2 e^{-x}$, où C_1 et C_2 sont des constantes arbitraires.

Montrons en conclusion comment on peut trouver la solution générale de l'équation (2) si l'on connaît une seule solution particulière.

Soit $y_1(x)$ une solution particulière de l'équation (2). Posons $y = y_1 z$. Alors $y' = y_1' z + y_1 z'$, $y'' = y_1'' z + 2y_1' z' + y_1 z''$. En portant les expressions de y, y' et y'' dans l'équation (2) et en réduisant les termes semblables, on obtient

$$[y_1'' + p(x) y_1' + q(x) y_1] z + z' [2y_1' + p(x) y_1] + y_1 z'' = 0.$$

Comme $y_1(x)$ est solution de l'équation (2), les premiers crochets sont nuls et l'équation prend la forme

$$z' [2y_1' + p(x) y_1] + y_1 z'' = 0.$$

On peut abaisser l'ordre de cette équation en posant $z' = u(x)$, où $u(x)$ est la nouvelle fonction à déterminer

$$u[2y_1' + p(x)\,y_1] + u'y_1 = 0.$$

On reconnaît ici une équation du premier ordre en u à variables séparables. Son intégration nous donne

$$\int \frac{du}{u} = -2\int \frac{dy_1}{y_1} - \int p(x)\,dx, \quad \ln|u| = -2\ln|y_1| - \int p(x)\,dx + \ln|C|$$

$$u = \pm \frac{C}{y_1^2} e^{-\int p(x)dx} \quad \text{ou} \quad u = \frac{C_1}{y_1^2} e^{-\int p(x)dx},$$

où C_1 est une constante arbitraire. En retournant à la variable z et en multipliant l'expression de z par y_1, on obtient la solution générale de l'équation (2):

$$y = C_1 y_1 \int \frac{1}{y_1^2} e^{-\int p(x)dx}\,dx + C_2 y_1,$$

où C_1 et C_2 sont des constantes arbitraires.

On laisse au lecteur le soin de trouver à titre d'exemple la solution générale de l'équation $y'' - y = 0$ (cf. exemple 1) en admettant que l'une des solutions particulières est connue.

3. Equations différentielles linéaires du second ordre avec second membre. Considérons maintenant les propriétés fondamentales des solutions de l'équation différentielle linéaire (1)

$$y'' + p(x)\,y' + q(x)\,y = f(x).$$

On a le théorème suivant.

Théorème 15.7. *La solution générale de l'équation* (1) *est la somme d'une solution particulière de* (1) *et de la solution générale de l'équation sans second membre associée.*

Démonstration. Soient $\tilde{y}(x)$ une solution particulière de l'équation (1), $Y(x) = C_1 y_1(x) + C_2 y_2(x)$, la solution générale de l'équation sans second membre associée (2), où C_1 et C_2 sont des constantes arbitraires. Montrons que la fonction

$$y = \tilde{y}(x) + Y(x) \tag{10}$$

est solution de l'équation (1). On a $y' = \tilde{y}'(x) + Y'(x)$, $y'' = \tilde{y}''(x) + Y''(x)$ et en portant dans (1):

$$\tilde{y}''(x) + Y''(x) + p(x)[\tilde{y}'(x) + Y'(x)] + q(x)[\tilde{y}(x) + Y(x)] =$$
$$= [Y''(x) + p(x)\,Y'(x) + q(x)\,Y(x)] + [\tilde{y}''(x) + p(x)\,\tilde{y}'(x) + q(x)\,\tilde{y}(x)] = 0 + f(x) = f(x).$$

La fonction $y = \tilde{y}(x) + Y(x)$ est bien une solution de (1).

Montrons maintenant qu'elle est la solution générale de (1). Prenons à cet effet une solution y quelconque de (1) et considérons la différence $y - \tilde{y}(x)$. Cette différence est solution de l'équation sans second membre (2). En effet,

$$[y - \tilde{y}(x)]'' + p(x)[y - \tilde{y}(x)]' + q(x)[y - \tilde{y}(x)] =$$
$$= [y'' + p(x)y' + q(x)y] - [\tilde{y}''(x) + p(x)\tilde{y}'(x) + q(x)\tilde{y}(x)] =$$
$$= f(x) - f(x) = 0.$$

Ce qui exprime que la différence $y - \tilde{y}(x)$ peut être mise sous la forme

$$y - \tilde{y}(x) = C_1^0 y_1(x) + C_2^0 y_2(x), \quad \text{d'où} \quad y = \tilde{y}(x) + C_1^0 y_1(x) + C_2^0 y_2(x),$$

où C_1^0 et C_2^0 sont des valeurs bien définies des constantes C_1 et C_2. Ainsi toute solution y de l'équation (1) se déduit de la formule (10) par un choix convenable des constantes C_1 et C_2, ce qui exprime que la fonction (10) est la solution générale de l'équation (1). ■

Donc pour trouver la solution générale d'une équation différentielle linéaire avec second membre il faut trouver la solution générale de l'équation sans second membre associée et une solution particulière de l'équation avec second membre. La détermination d'une solution particulière est un problème ardu dans le cas général. Montrons comment on peut trouver une solution particulière d'une équation avec second membre par la méthode de variation des constantes si l'on connaît la solution générale de l'équation sans second membre associée.

Soit $Y = C_1 y_1(x) + C_2 y_2(x)$ la solution générale de l'équation sans second membre (2). Cherchons une solution particulière de (1) sous la forme

$$y = C_1(x)\, y_1(x) + C_2(x)\, y_2(x), \tag{11}$$

en traitant C_1 et C_2 comme des fonctions de x. Une dérivation de (11) nous donne

$$y' = C_1'(x)\, y_1(x) + C_1(x)\, y_1'(x) + C_2'(x)\, y_2(x) + C_2(x)\, y_2'(x). \tag{12}$$

Choisissons les fonctions $C_1(x)$ et $C_2(x)$ de telle sorte que

$$C_1'(x)\, y_1(x) + C_2'(x)\, y_2(x) = 0. \tag{13}$$

L'égalité (12) devient alors

$$y' = C_1(x)\, y_1'(x) + C_2(x)\, y_2'(x).$$

En dérivant, on obtient

$$y'' = C_1'(x)\,y_1'(x) + C_1(x)\,y_1''(x) + C_2'(x)\,y_2'(x) + C_2(x)\,y_2''(x).$$

En portant les expressions de y, y' et y'' dans (1) et en réduisant les termes semblables, on trouve

$$C_1(x)\,[y_1''(x) + p(x)\,y_1'(x) + q(x)\,y_1(x)] + \\ + C_2(x)\,[y_2''(x) + p(x)\,y_2'(x) + q(x)\,y_2(x)] + \\ + C_1'(x)\,y_1'(x) + C_2'(x)\,y_2'(x) = f(x).$$

Les crochets sont tous deux nuls, puisque $y_1(x)$ et $y_2(x)$ sont solutions de l'équation sans second membre. Donc cette égalité devient

$$C_1'(x)\,y_1'(x) + C_2'(x)\,y_2'(x) = f(x). \tag{14}$$

La fonction (11) est donc solution de l'équation (1) si les fonctions $C_1(x)$ et $C_2(x)$ satisfont aux équations (13) et (14). En les regroupant, on obtient le système d'équations

$$\begin{cases} C_1'(x)\,y_1(x) + C_2'(x)\,y_2(x) = 0, \\ C_1'(x)\,y_1'(x) + C_2'(x)\,y_2'(x) = f(x), \end{cases} \tag{15}$$

où $C_1'(x)$ et $C_2'(x)$ sont des fonctions inconnues, $y_1(x)$, $y_2(x)$ $y_1'(x)$, $y_2'(x)$ et $f(x)$, des fonctions connues. Le déterminant de ce système est le wronskien

$$W(x) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix}$$

des solutions linéairement indépendantes $y_1(x)$ et $y_2(x)$ de l'équation sans second membre (2). Comme $W(x) \neq 0$ d'après le théorème 15.5, le système (15) admet une solution unique

$$C_1'(x) = \varphi_1(x), \quad C_2'(x) = \varphi_2(x),$$

où $\varphi_1(x)$ et $\varphi_2(x)$ sont des fonctions connues dont l'intégration nous donne $C_1(x)$ et $C_2(x)$. En portant les expressions de $C_1(x)$ et $C_2(x)$ dans (11), on obtient la solution particulière cherchée de l'équation (1).

Exemple 2. Trouver une solution particulière de l'équation $y'' - y = x$.

Solution. La solution générale de l'équation sans second membre associée est $Y(x) = C_1e^x + C_2e^{-x}$ (cf. exemple 1 du n° 2). On cherchera donc une solution particulière de l'équation complète sous la forme

$$\tilde{y}(x) = C_1(x)\,e^x + C_2(x)\,e^{-x}. \tag{16}$$

Le système (15) en $C_1'(x)$ et $C_2'(x)$ s'écrit ici

$$\begin{cases} C_1'(x)\,e^x + C_2'(x)\,e^{-x} = 0, \\ C_1'(x)\,e^x - C_2'(x)\,e^{-x} = x. \end{cases}$$

En ajoutant ces équations on trouve que $C_1'(x) = \frac{1}{2}\,xe^{-x}$. D'où, en intégrant,

$$C_1(x) = -\frac{1}{2}\,(x+1)\,e^{-x}.$$

On omet d'écrire la constante arbitraire, puisqu'on cherche une solution particulière quelconque. En portant $C_1'(x)$ dans la première équation du système, on trouve $C_2'(x) = -\frac{1}{2}\,xe^x$, d'où, en intégrant,

$$C_2(x) = -\frac{1}{2}\,(x-1)\,e^x.$$

La substitution de $C_1(x)$ et $C_2(x)$ dans (16) nous donne une solution particulière $\tilde{y}$ de l'équation complète

$$\tilde{y}(x) = \left[-\frac{1}{2}\,(x+1)\,e^{-x}\right]e^x + \left[-\frac{1}{2}\,(x-1)\,e^x\right]e^{-x} = -x.$$

Remarquons qu'en vertu de (10) la solution générale de l'équation complète est

$$y = \tilde{y}(x) + Y(x) = -x + C_1e^x + C_2e^{-x},$$

où C_1 et C_2 sont des constantes arbitraires.

§ 4. Equations différentielles linéaires du second ordre à coefficients constants

Considérons un cas particulier important d'équations différentielles linéaires du second ordre, le cas où les fonctions $p(x)$ et $q(x)$ sont des constantes. De telles équations sont dites équations linéaires à coefficients constants.

1. Equations différentielles linéaires du second ordre sans second membre à coefficients constants. Soit une équation linéaire du second ordre

$$y'' + py' + qy = 0, \tag{1}$$

où p et q sont des nombres réels.

Théorème 15.8. 1) *Si k est une racine réelle de l'équation*

$$k^2 + pk + q = 0, \tag{2}$$

la fonction $y = e^{kx}$ est solution de l'équation (1).

2) *Si* $k_1 = \alpha + i\beta$ *et* $k_2 = \alpha - i\beta$ ($\beta \neq 0$) *sont des racines complexes de l'équation* (2), *les fonctions* $y_1 = e^{\alpha x}\cos\beta x$ *et* $y_2 = e^{\alpha x}\sin\beta x$ *sont solutions de l'équation* (1).

Démonstration. 1) Soit $y = e^{kx}$ ($k = \text{const}$). Alors $y' = ke^{kx}$, $y'' = k^2e^{kx}$. En portant y, y' et y'' dans l'équation (1), on obtient

$$e^{kx}(k^2 + pk + q) = 0.$$

Comme $e^{kx} \neq 0$, en simplifiant par e^{kx}, on trouve

$$k^2 + pk + q = 0.$$

Donc si k est une racine de l'équation (2), la fonction $y = e^{kx}$ est solution de l'équation (1).

2) Comme dans le cas 1), on vérifie sans peine que les fonctions $y_1 = e^{\alpha x}\cos\beta x$ et $y_2 = e^{\alpha x}\sin\beta x$ sont solutions de l'équation (1). ■

L'équation (2) s'appelle *équation caractéristique* de l'équation (1).

L'équation caractéristique (2) est un trinôme du second degré qui par conséquent a deux racines désignées par k_1 et k_2.

Théorème 15.9. 1) *Si les racines de l'équation caractéristique sont réelles et distinctes* ($k_1 \neq k_2$), *la solution générale de l'équation* (1) *est de la forme*

$$y = C_1e^{k_1x} + C_2e^{k_2x};$$

2) *si les racines de l'équation caractéristique sont réelles et égales* ($k_1 = k_2$), *la solution générale est*

$$y = C_1e^{k_1x} + C_2xe^{k_1x};$$

3) *si les racines de l'équation caractéristique sont complexes* ($k_{1,2} = \alpha \pm i\beta$. $\beta \neq 0$), *la solution générale est de la forme*

$$y = e^{\alpha x}(C_1\cos\beta x + C_2\sin\beta x).$$

Démonstration. 1) Supposons que les racines k_1 et k_2 sont distinctes. D'après le théorème 15.8, les fonctions $y_1 = e^{k_1x}$ et $y_2 = e^{k_2x}$ sont solutions particulières de l'équation (1). Ces solutions sont linéairement indépendantes, puisque $\frac{y_2}{y_1} = e^{(k_2-k_1)x} \neq \text{const}$ ($k_1 \neq k_2$). Donc d'après le théorème 15.6, la solution générale de l'équation (1) est

$$y = C_1e^{k_1x} + C_2e^{k_2x}.$$

2) Supposons que les racines k_1 et k_2 sont égales. D'après le théorème 15.8, la fonction $y_1 = e^{k_1x}$ est solution particulière de l'équation (1). Trouvons une autre solution particulière linéairement indépendante de la première. Cherchons-la sous la forme

$y_2 = z(x)\, e^{k_1 x}$, où $z(x)$ est la nouvelle fonction inconnue. Une dérivation nous donne

$$y_2' = z' e^{k_1 x} + k_1 z e^{k_1 x} = e^{k_1 x}(z' + k_1 z),$$

$$y_2'' = z'' e^{k_1 x} + 2k_1 z' e^{k_1 x} + k_1^2 z e^{k_1 x} = e^{k_1 x}(z'' + 2k_1 z' + k_1^2 z).$$

En portant y_2, y_2' et y_2'' dans le premier membre de l'équation (1), on obtient

$$e^{k_1 x}\left[z'' + 2\left(k_1 + \frac{p}{2}\right) z' + (k_1^2 + pk_1 + q) z\right] = 0.$$

Par hypothèse, $k_1^2 + pk_1 + q = 0$. Par ailleurs, $k_1 = k_2 = -p/2$, donc $k_1 + p/2 = 0$. Et, par suite, pour trouver la fonction $z(x)$ il faut intégrer l'équation $z'' = 0$. Une double intégration nous donne $z' = \widetilde{C}_1$, $z = \widetilde{C}_1 x + \widetilde{C}_2$, où $\widetilde{C}_1$ et $\widetilde{C}_2$ sont des constantes arbitraires. En posant $\widetilde{C}_1 = 1$ et $\widetilde{C}_2 = 0$, on trouve $z(x) = x$. Donc $y_2 = xe^{k_1 x}$ est une autre solution particulière de l'équation (1). Les solutions y_1 et y_2 sont linéairement indépendantes, puisque $y_1/y_2 = 1/x \neq \text{const}$ et d'après le théorème 15.6, la solution générale de l'équation (1) est

$$y = C_1 e^{k_1 x} + C_2 x e^{k_1 x}.$$

3) Supposons que les racines k_1 et k_2 sont complexes, c'est-à-dire que $k_{1,2} = \alpha \pm i\beta$ ($\beta \neq 0$). D'après le théorème 15.8, les fonctions $y_1 = e^{\alpha x} \cos \beta x$ et $y_2 = e^{\alpha x} \sin \beta x$ sont solutions particulières de l'équation (1). Ces solutions sont linéairement indépendantes, puisque $y_1/y_2 = \operatorname{cotg} \beta x \neq \text{const}$. Donc la solution générale de l'équation (1) est

$$y = e^{\alpha x}(C_1 \cos \beta x + C_2 \sin \beta x). \blacksquare$$

Exemple 1. Trouver la solution générale de l'équation $y'' + y' - 2y = 0$.

Solution. L'équation caractéristique est $k^2 + k - 2 = 0$. Ses racines $k_1 = 1$ et $k_2 = -2$ sont réelles et distinctes. Donc les solutions particulières correspondantes sont $y_1 = e^x$, $y_2 = e^{-2x}$ et la solution générale est $y = C_1 e^x + C_2 e^{-2x}$.

Exemple 2. Trouver la solution générale de l'équation $y'' - 2y' + y = 0$.

Solution. L'équation caractéristique est de la forme $k^2 - 2k + 1 = 0$. Ses racines $k_1 = k_2 = 1$ sont réelles et confondues. Donc les solutions particulières correspondantes sont $y_1 = e^x$, $y_2 = xe^x$ et la solution générale est $y = C_1 e^x + C_2 x e^x$.

Exemple 3. Trouver la solution générale de l'équation $y'' - 4y' + 13y = 0$.

Solution. L'équation caractéristique est $k^2 - 4k + 13 = 0$. Ses racines $k_{1,2} = 2 \pm 3i$ sont complexes. Les solutions particulières correspondantes sont $y_1 = e^{2x} \cos 3x$, $y_2 = e^{2x} \sin 3x$. La solution générale est donc $y = e^{2x} (C_1 \cos 3x + C_2 \sin 3x)$.

2. Equations différentielles linéaires avec second membre du second ordre à coefficients constants. Considérons l'équation linéaire avec second membre du second ordre

$$y'' + py' + qy = f(x), \tag{3}$$

où p et q sont des nombres réels; $f(x)$, une fonction continue.

On sait que la solution générale de cette équation est la somme d'une solution particulière de l'équation complète et de la solution générale de l'équation sans second membre associée. Nous savons déterminer la solution générale de l'équation sans second membre associée. Reste donc à envisager le problème de la détermination d'une solution particulière. On pourrait certes se servir de la méthode de variation des constantes. Mais si le second membre de l'équation (3) est un polynôme, ou une exponentielle, ou une fonction circulaire $\sin \beta x$ ou $\cos \beta x$, ou combinaison linéaire des fonctions énumérées, alors on peut trouver une solution particulière par la méthode des coefficients indéterminés, méthode qui n'implique pas d'intégration.

Voyons les divers types de second membre de l'équation (3).

1) $f(x) = P_n(x)$, où $P_n(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ est un polynôme de degré n. Dans ce cas il faut chercher une solution particulière $\tilde{y}$ sous la forme

$$\tilde{y} = Q_n(x)\, x^r,$$

où $Q_n(x)$ est un polynôme de même degré que $P_n(x)$, et r la multiplicité de la racine 0 de l'équation caractéristique.

Exemple 4. Trouver la solution générale de l'équation $y'' - 2y' + y = x + 1$.

Solution. La solution générale de l'équation sans second membre associée est $Y = (C_1 + C_2 x)e^x$ (cf. exemple 2). Puisque le second membre est un binôme du premier degré et qu'aucune racine de l'équation caractéristique $k^2 - 2k + 1 = 0$ n'est nulle ($k_2 = k_1 = 1$), on cherchera une solution particulière sous la forme

$$\tilde{y} = (Ax + B)\, x^0 = Ax + B,$$

où A et B sont des coefficients inconnus. En dérivant deux fois $\tilde{y} = Ax + B$ et en substituant $\tilde{y}$, $\tilde{y}'$ et $\tilde{y}''$ dans l'équation, on

obtient

$$-2A + Ax + B = x + 1.$$

L'identification des coefficients de x nous donne $A = 1$, $-2A + B = 1$, d'où $A = 1$, $B = 3$. Donc une solution particulière de cette équation est $\tilde{y} = x + 3$, et la solution générale

$$y = (C_1 + C_2 x)\, e^x + (x + 3).$$

2) $f(x) = e^{\alpha x} P_n(x)$, où $P_n(x)$ est un polynôme de degré n. On cherchera une solution particulière y sous la forme

$$\tilde{y} = Q_n(x)\, x^r e^{\alpha x},$$

où $Q_n(x)$ est un polynôme du même degré que $P_n(x)$, et r la multiplicité de la racine α de l'équation caractéristique. Si $\alpha = 0$, $f(x) = P_n(x)$ et l'on retrouve le cas 1).

Exemple 5. Trouver la solution générale de l'équation $y'' - 4y' + 3y = xe^x$.

Solution. Les racines de l'équation caractéristique $k^2 - 4k + 3 = 0$ sont $k_1 = 1$, $k_2 = 3$. La solution générale de l'équation sans second membre associée est alors $Y = C_1 e^x + C_2 e^{3x}$. Le second membre est le produit d'un monôme du premier degré par la fonction exponentielle $e^{\alpha x}$ pour $\alpha = 1$. Ici $r = 1$ car une seule racine de l'équation caractéristique, la racine k_1, est égale à 1 ; d'autre part, $P_n(x) = x$ est un monôme du premier degré. On cherchera donc une solution particulière de cette équation sous la forme

$$\tilde{y} = (Ax + B)\, xe^x = (Ax^2 + Bx)\, e^x.$$

En dérivant et en substituant dans l'équation, on obtient

$$-4Ax + 2A - 2B = x.$$

Une identification des coefficients nous conduit au système $-4A = 1$, $2A - 2B = 0$, d'où $A = -\frac{1}{4}$, $B = -\frac{1}{4}$. En portant les valeurs de A et B dans l'expression de y, on trouve une solution particulière $\tilde{y} = -\frac{1}{4}(x^2 + x)\, e^x$ et la solution générale est

$$y = \tilde{y} + Y = C_1 e^x + C_2 e^{3x} - \frac{1}{4}(x^2 + x)\, e^x.$$

3) $f(x) = a \cos \beta x + b \sin \beta x$, où a, b et β sont des nombres connus. Il faut chercher une solution particulière sous la forme

$$\tilde{y} = (A \cos \beta x + B \sin \beta x)\, x^r,$$

où A et B sont des coefficients inconnus, r la multiplicité de la racine $i\beta$ de l'équation caractéristique.

Exemple 6. Trouver la solution générale de l'équation $y'' + y = \sin x$.

Solution. L'équation caractéristique $k^2 + 1 = 0$ admet les racines $k_{1,2} = \pm i$. La solution générale de l'équation sans second membre associée est $Y = C_1 \cos x + C_2 \sin x$. Au second membre on a une fonction circulaire $\sin x$, c'est-à-dire que $a = 0$, $b = 1$, $\beta = 1$. Puisque $i\beta = i$ est une racine de l'équation caractéristique, on a $r = 1$ et on cherchera une solution particulière sous la forme

$$\tilde{y} = (A \cos x + B \sin x)\, x.$$

En dérivant et en substituant dans l'équation, on trouve

$$2\,(-A \sin x + B \cos x) = \sin x,$$

d'où $A = -1/2$, $B = 0$. Donc $\tilde{y} = -\frac{1}{2} x \cos x$ et la solution générale est

$$y = \tilde{y} + Y = C_1 \cos x + C_2 \sin x - \frac{1}{2} x \cos x.$$

Exemple 7. Trouver la solution générale de l'équation $y'' + y = \sin 2x$.

Solution. Cette équation diffère de la précédente par le fait que $\beta = 2$. Comme $i\beta = 2i$ n'est pas une racine de l'équation caractéristique, on a $r = 0$ et il faut chercher une solution particulière sous la forme

$$\tilde{y} = A \cos 2x + B \sin 2x.$$

En dérivant et en substituant dans l'équation, on trouve

$$-3A \cos 2x - 3B \sin 2x = \sin 2x,$$

d'où $A = 0$, $B = -1/3$, c'est-à-dire qu'une solution particulière est $\tilde{y} = -\frac{1}{3} \sin 2x$ et la solution générale est

$$y = \tilde{y} + Y = C_1 \cos x + C_2 \sin x - \frac{1}{3} \sin 2x.$$

4) $f(x) = e^{\alpha x}\,[P_n(x) \cos \beta x + P_m(x) \sin \beta x]$, où $P_n(x)$ est un polynôme de degré n et $P_m(x)$ un polynôme de degré m. On cherchera une solution particulière sous la forme

$$\tilde{y} = x^r e^{\alpha x}\,[Q_1(x) \cos \beta x + Q_2(x) \sin \beta x],$$

où $Q_1(x)$ et $Q_2(x)$ sont des polynômes de degré $s = \max\{n, m\}$, r le nombre de racines de l'équation caractéristique égales à $\alpha + i\beta$.

Exemple 8. Trouver la solution générale de l'équation $y'' - y = 3e^{2x} \cos x$.

Solution. Les racines de l'équation caractéristique $k^2 - 1 = 0$ sont $k_{1,2} = \pm 1$. La solution générale de l'équation sans second membre associée est $Y = C_1 e^x + C_2 e^{-x}$. Le second membre de l'équation est le produit d'un polynôme de degré nul, d'une exponentielle et d'une fonction circulaire, donc $P_n(x) = 3$, $P_m(x) = 0$, $s = 0$. Le nombre $\alpha + i\beta = 2 + i$ n'est pas racine de l'équation caractéristique, donc $r = 0$ et il faut chercher une solution particulière sous la forme

$$\tilde{y} = e^{2x} (A \cos x + B \sin x).$$

En dérivant et en portant dans l'équation, on obtient

$$(2A + 4B) \cos x + (2B - 4A) \sin x = 3 \cos x.$$

Une identification des coefficients de $\cos x$ et $\sin x$ nous donne

$$2A + 4B = 3, \quad -4A + 2B = 0,$$

d'où $A = 3/10$, $B = 3/5$. Donc une solution particulière est $\tilde{y} = e^{2x} \left(\frac{3}{10} \cos x + \frac{3}{5} \sin x \right)$, et la solution générale est

$$y = \tilde{y} + Y = e^{2x} \left(\frac{3}{10} \cos x + \frac{3}{5} \sin x \right) + C_1 e^x + C_2 e^{-x}.$$

Exemple 9. Sachant les racines de l'équation caractéristique et le second membre $f(x)$, trouver une solution particulière $\tilde{y}$ de l'équation complète :

a) $k_1 = 3 + i2$, $k_2 = 3 - i2$, $f(x) = 8e^{3x} \sin 2x$;

b) $k_1 = k_2 = -3$, $f(x) = 2xe^{-3x} \sin x$;

c) $k_1 = 1$, $k_2 = -3$, $f(x) = e^x (1 - x) \cos 3x$;

d) $k_1 = 1 + i2$, $k_2 = 1 - i2$, $f(x) = e^x (\cos 2x - 3 \sin 2x)$;

e) $k_1 = 2 + i\frac{1}{2}$, $k_2 = 2 - i\frac{1}{2}$,

$$f(x) = e^{2x} \left[(x^3 - 1) \cos \frac{x}{2} + x \sin \frac{x}{2} \right].$$

Solution. a) On a $\alpha = 3$, $\beta = 2$, $P_n(x) = 0$, $P_m(x) = 8$, $s = 0$. Le nombre $\alpha + i\beta = 3 + 2i$ étant racine de l'équation caractéristique, il vient $r = 1$. Donc $\tilde{y} = xe^{3x} (A \cos 2x + B \sin 2x)$.

b) On a $\alpha = -3$, $\beta = 1$, $P_n(x) = 0$, $P_m(x) = 2x$, $m = 1$, $s = 1$. Le nombre $\alpha + i\beta = -3 + i$ n'étant pas une racine de l'équation caractéristique, il vient $r = 0$. Donc $\tilde{y} = e^{-3x} [(Ax + B) \cos x + (Cx + D) \sin x]$.

c) On a $\alpha = 1$, $\beta = 3$, $P_n(x) = 1 - x$, $P_m(x) = 0$, $n = 1$, $s = 1$. Vu que $\alpha + i\beta = 1 + 3i$ n'est pas une racine de l'équa-

tion caractéristique, $r = 0$. Donc $\tilde{y} = e^x [(Ax + B) \cos 3x + (Cx + D) \sin 3x]$.

d) On a $\alpha = 1$, $\beta = 2$, $P_n(x) = 1$, $P_m(x) = -3$, $s = 0$. Le nombre $\alpha + i\beta = 1 + 2i$ étant racine de l'équation caractéristique, $r = 1$. Donc $\tilde{y} = xe^x (A \cos 2x + B \sin 2x)$.

e) On a $\alpha = 2$, $\beta = \frac{1}{2}$, $P_n(x) = (x^3 - 1)$, $P_m(x) = x$, $n = 3$, $m = 1$, $s = 3$. Le nombre $\alpha + i\beta = 2 + \frac{1}{2}i$ étant racine de l'équation caractéristique, $r = 1$. Donc $\tilde{y} = xe^{2x} \left[(Ax^3 + Bx^2 + Cx + D) \cos \frac{x}{2} + (Ex^3 + Fx^2 + Gx + H) \sin \frac{x}{2}\right]$.

Prouvons en conclusion un théorème qui est largement appliqué à l'intégration des équations différentielles linéaires dont le second membre est la somme de plusieurs termes.

Théorème 15.10. *Si $\tilde{y}_1$ est une solution particulière de l'équation*

$$y'' + py' + qy = f_1(x), \tag{4}$$

et $\tilde{y}_2$, une solution particulière de l'équation

$$y'' + py' + qy = f_2(x), \tag{5}$$

la somme $\tilde{y}_1 + \tilde{y}_2$ est une solution particulière de l'équation

$$y'' + py' + qy = f_1(x) + f_2(x). \tag{6}$$

Démonstration. Composons la somme $\tilde{y}_1 + \tilde{y}_2$ et portons-la dans le premier membre de l'équation (6). On a

$$(\tilde{y}_1 + \tilde{y}_2)'' + p(\tilde{y}_1 + \tilde{y}_2)' + q(\tilde{y}_1 + \tilde{y}_2) =$$

$$= (\tilde{y}_1'' + p\tilde{y}_1' + q\tilde{y}_1) + (\tilde{y}_2'' + p\tilde{y}_2' + q\tilde{y}_2) = f_1 + f_2,$$

donc $\tilde{y}_1 + \tilde{y}_2$ est une solution particulière de l'équation (6). ■

Exemple 10. Trouver la solution générale de l'équation $y'' - 2y' + y = \sin x + e^{-x}$.

Solution. Les racines de l'équation caractéristique $k^2 - 2k + 1 = 0$ sont $k_{1,2} = 1$, donc la solution générale de l'équation sans second membre associée est $Y = C_1 e^x + C_2 x e^x = e^x (C_1 + C_2 x)$.

Le second membre de l'équation étant la somme des fonctions $\sin x$ et e^{-x}, on cherchera en vertu du théorème 15.10 une solution particulière sous la forme

$$\tilde{y} = \tilde{y}_1 + \tilde{y}_2,$$

où $\tilde{y}_1$ est une solution particulière de l'équation $y'' - 2y' + y = \sin x$ et $\tilde{y}_2$, une solution particulière de l'équation $y'' - 2y' + y = e^{-x}$.

Trouvons d'abord la solution particulière $\tilde{y}_1$. Vu que le nombre $i\beta = i$ n'est pas une racine de l'équation caractéristique ($r = 0$), on cherchera la solution particulière $\tilde{y}_1$ sous la forme $\tilde{y}_1 = A \sin x + B \cos x$. En portant $\tilde{y}_1$, $\tilde{y}_1'$ et $\tilde{y}_1''$ dans l'équation $y'' - 2y' + y = \sin x$ et en identifiant les coefficients de $\sin x$ et $\cos x$, on obtient $-2A = 0$, $2B = 1$, d'où $A = 0$, $B = 1/2$ et par suite $\tilde{y}_1 = \frac{1}{2} \cos x$.

Trouvons maintenant la solution particulière $\tilde{y}_2$. On la cherchera sous la forme $\tilde{y}_2 = Ae^{-x}$, puisque $\alpha = -1$ n'est pas une racine de l'équation caractéristique. En portant $\tilde{y}_2$, $\tilde{y}_2'$ et $\tilde{y}_2''$ dans l'équation $y'' - 2y' + y = e^{-x}$, on trouve $A = \frac{1}{4}$. Donc $\tilde{y}_2 = \frac{1}{4} e^{-x}$.

Donc une solution particulière est

$$\tilde{y} = \tilde{y}_1 + \tilde{y}_2 = \frac{1}{2} \cos x + \frac{1}{4} e^{-x}$$

et la solution générale,

$$y = \tilde{y} + Y = \frac{1}{2} \cos x + \frac{1}{4} e^{-x} + C_1 e^x + C_2 x e^x.$$

§ 5. Application des équations différentielles linéaires à l'étude des phénomènes vibratoires

Les équations différentielles linéaires constituent un puissant outil d'étude des phénomènes vibratoires qui occupent une place importante en physique et dans les techniques actuelles. Etudions le mouvement oscillatoire d'un solide suspendu à un ressort vertical.

Position du problème. Soit un solide de masse m suspendu à l'extrémité d'un ressort et supposons que les seuls mouvements possibles du solide soient des translations verticales. Si l'on comprime ou étire le ressort avec le solide, il prend un mouvement oscillatoire. Trouvons l'équation de ce mouvement, c'est-à-dire la formule qui exprime l'écart du solide par rapport à sa position d'équilibre à tout instant t.

Faisons coïncider l'origine des coordonnées avec la position d'équilibre du solide et orientons l'axe Oy verticalement vers le

haut. Désignons par l_0 la distance de l'extrémité du ressort au repos sans le solide à la position d'équilibre du solide, et par y l'écart du solide par rapport à la position d'équilibre à l'instant t (fig. 227). Le solide est soumis à l'action de la résultante des trois forces suivantes: 1) le poids mg qui est dirigé verticalement vers le bas; 2) la résistance de l'air opposée au mouvement du solide et proportionnelle à la vitesse, c'est-à-dire une force égale à $\lambda \frac{dy}{dx}$, où λ est le coefficient de proportionnalité; 3) la force élastique du ressort dirigée verticalement vers le haut et proportionnelle, en vertu de la loi de Hooke *), à l'allongement du ressort, c'est-à-dire qu'elle est égale à $c(l_0 - y)$, où c est un coefficient de proportionnalité appelé *coefficient de rigidité du ressort* (la masse du ressort est négligeable).

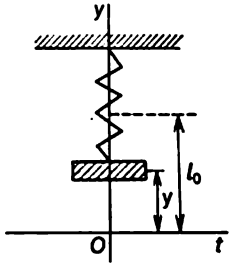


Fig. 227

D'après la deuxième loi de Newton, l'équation du mouvement du solide est

$$m \frac{d^2y}{dt^2} = c(l_0 - y) - mg - \lambda \frac{dy}{dt}.$$

Puisque dans la position d'équilibre ($y = 0$) le poids mg est compensé par la force élastique du ressort, on a $mg = cl_0$. Donc

$$m \frac{d^2y}{dt^2} = -cy - \lambda \frac{dy}{dt}$$

ou

$$m \frac{d^2y}{dt^2} + \lambda \frac{dy}{dt} + cy = 0. \tag{1}$$

L'équation différentielle obtenue s'appelle *équation des oscillations libres d'un solide suspendu à un ressort.*

Si le solide est soumis à l'action d'une force extérieure dirigée verticalement vers le haut, dont l'intensité $F(t)$ dépend du temps t, l'équation (1) devient

$$m \frac{d^2y}{dt^2} + \lambda \frac{dy}{dt} + cy = F(t). \tag{2}$$

L'équation (2) s'appelle *équation des oscillations forcées d'un solide suspendu à un ressort.*

En divisant l'équation (2) par m et en posant $\frac{c}{m} = \omega^2$, $\frac{\lambda}{m} = 2\mu$, $\frac{F(t)}{m} = f(t)$, on obtient la forme définitive de l'équa-

*) Hooke Robert (1635-1703), physicien anglais.

tion des oscillations forcées :

$$\frac{d^2y}{dt^2} + 2\mu \frac{dy}{dt} + \omega^2 y = f(t). \tag{3}$$

L'équation (3) est une *équation différentielle linéaire avec second membre du second ordre à coefficients constants.*

Etudions maintenant les oscillations à l'aide des solutions des équations linéaires à coefficients constants.

Oscillations libres. Supposons que le solide n'est pas soumis à l'action de la force extérieure $f(t)$ et que la résistance est nulle, i.e. $\mu = 0$. L'équation (3) devient alors

$$\frac{d^2y}{dt^2} + \omega^2 y = 0.$$

On reconnaît ici une équation linéaire sans second membre. Les racines de l'équation caractéristique $k^2 + \omega^2 = 0$ étant $k_{1,2} = \pm i\omega$, la solution générale est

$$y = C_1 \cos \omega t + C_2 \sin \omega t,$$

où C_1 et C_2 sont des constantes arbitraires. Pour alléger les raisonnements remplaçons les constantes arbitraires C_1 et C_2 par des constantes $A > 0$ et φ en posant $C_1 = A \sin \varphi$ et $C_2 = A \cos \varphi$ (d'où $A = \sqrt{C_1^2 + C_2^2}$ et $\operatorname{tg} \varphi = C_1/C_2$). Alors

$$y = A \sin \varphi \cos \omega t + A \cos \varphi \sin \omega t,$$

et la solution générale peut être mise sous la forme

$$y = A \sin (\omega t + \varphi). \tag{4}$$

La formule (4) est l'équation du mouvement d'un solide suspendu à un ressort. D'après cette formule le solide effectue des oscillations harmoniques autour de la position d'équilibre. La quantité A s'appelle *amplitude des oscillations*, ω, *pulsation* et φ, *phase initiale.*

Pour déterminer une solution particulière, il faut se donner des conditions initiales du mouvement. Supposons que l'écart et la vitesse du solide sont connus à l'instant initial $t = 0$:

$$y|_{t=0} = y_0, \quad y'|_{t=0} = y_0'. \tag{5}$$

Une dérivation nous donne

$$y' = A\omega \cos (\omega t + \varphi); \tag{6}$$

en portant les conditions initiales (5) dans (4) et (6), on obtient

$$\begin{cases} y_0 = A \sin \varphi, \\ y_0' = A\omega \cos \varphi. \end{cases} \tag{7}$$

En exprimant les constantes arbitraires A et φ en fonction de ω, y_0 et y'_0 et en les portant dans (4), on trouve la solution particulière qui satisfait aux conditions initiales (5).

Des formules (7) il s'ensuit en particulier que les constantes A et φ dépendent de la pulsation ω et des conditions initiales du mouvement. La pulsation, quant à elle, ne dépend pas de ces conditions, mais du rapport du coefficient de rigidité à la masse du solide : $\omega = \sqrt{c/m}$.

Supposons maintenant que $f(t) = 0$ et $\mu \neq 0$. L'équation (3) devient alors

$$\frac{d^2y}{dt^2} + 2\mu \frac{dy}{dt} + \omega^2 y = 0. \tag{8}$$

Les racines de l'équation caractéristique $k^2 + 2\mu k + \omega^2 = 0$ sont $k_{1,2} = -\mu \pm \sqrt{\mu^2 - \omega^2}$. On distinguera trois cas.

1) $\mu > \omega$. Les racines $k_{1,2} = -\mu \pm \sqrt{\mu^2 - \omega^2}$ sont réelles, distinctes et strictement négatives. La solution générale de l'équation (8) est donc

$$y = C_1 e^{(-\mu + \sqrt{\mu^2 - \omega^2})t} + C_2 e^{(-\mu - \sqrt{\mu^2 - \omega^2})t}.$$

On voit sur cette formule que le solide n'effectue pas d'oscillations, et que l'écart y tend vers 0, c'est-à-dire que le solide tend vers la position d'équilibre lorsque $t \to +\infty$. On dit alors que le solide accomplit un mouvement amorti non périodique.

2) $\mu = \omega$. Les racines $k_{1,2} = -\mu$ sont réelles, égales et strictement négatives. La solution générale de l'équation (8) est alors

$$y = C_1 e^{-\mu t} + C_2 t e^{-\mu t} = e^{-\mu t}(C_1 + C_2 t).$$

Le solide effectue un mouvement identique au précédent.

3) $\mu < \omega$. Les racines $k_{1,2} = -\mu \pm i\sqrt{\omega^2 - \mu^2}$ sont conjuguées complexes. La solution générale de l'équation (8) est

$$y = e^{-\mu t}(C_1 \cos \tilde{\omega} t + C_2 \sin \tilde{\omega} t),$$

où $\tilde{\omega} = \sqrt{\omega^2 - \mu^2}$. Si l'on remplace C_1 et C_2 par les constantes A et φ, cette solution devient

$$y = Ae^{-\mu t} \sin(\tilde{\omega} t + \varphi).$$

Contrairement au cas de la formule (4), l'amplitude dépend ici du temps t. Puisque $\mu < 0$, cette amplitude tend vers 0 lorsque $t \to +\infty$. Donc le solide accomplit des oscillations amorties libres autour de la position d'équilibre.

Oscillations forcées. Résonance. Traitons maintenant le cas où le système oscillatoire est soumis à l'action d'une force exté-

rieure périodique $f(t) = a \sin \omega_1 t$ en supposant pour simplifier que la résistance de l'air est négligeable ($\mu = 0$). L'équation (8) devient alors

$$\frac{d^2y}{dt^2} + \omega^2 y = a \sin \omega_1 t. \tag{9}$$

C'est une équation linéaire avec second membre à coefficients constants. On sait que la solution générale de cette équation est la somme de la solution générale Y de l'équation sans second membre associée [cf. formule (4)] et d'une solution particulière $\tilde{y}$ de l'équation complète.

On distinguera deux cas.

a) $\omega \neq \omega_1$, c'est-à-dire que la pulsation de la force périodique extérieure est distincte de celle des oscillations libres du solide. Puisque $i\omega_1$ n'est pas confondu avec une racine de l'équation caractéristique $k^2 + \omega^2 = 0$, on cherchera une solution particulière sous la forme

$$\tilde{y} = A \cos \omega_1 t + B \sin \omega_1 t.$$

En dérivant deux fois $\tilde{y}$, et en substituant $\tilde{y}$ et $\tilde{y}''$ dans (9), on trouve $B = \frac{a}{\omega^2 - \omega_1^2}$, $A = 0$. Donc

$$\tilde{y} = \frac{a}{\omega^2 - \omega_1^2} \sin \omega_1 t$$

et la solution générale de l'équation (9) est

$$y = \tilde{y} + Y = \frac{a}{\omega^2 - \omega_1^2} \sin \omega_1 t + A \sin (\omega t + \varphi). \tag{10}$$

De la formule (10) il résulte que la solution particulière $\tilde{y}$ décrit les oscillations engendrées par la force extérieure, la solution générale $Y = A \sin (\omega t + \varphi)$, les oscillations libres du solide, et la solution générale y, un mouvement oscillatoire qui est la composition de deux mouvements oscillatoires de pulsations distinctes ω et ω_1.

Dans ce cas l'amplitude est constante et si ω et ω_1 sont voisines, le solide effectue des oscillations d'une grande amplitude autour de la position d'équilibre.

b) $\omega = \omega_1$, c'est-à-dire que la pulsation de la force périodique extérieure est confondue avec celle des oscillations libres du solide. Puisque $i\omega_1 = i\omega$ est une racine de l'équation caractéristique $k^2 + \omega^2 = 0$, on cherchera une solution particulière sous la forme

$$\tilde{y} = t (A \cos \omega t + B \sin \omega t).$$

En dérivant $\tilde{y}$ deux fois et en portant $\tilde{y}$ et $\tilde{y}''$ dans l'équation (9), on trouve $A = -\frac{a}{2\omega}$, $B = 0$. Donc

$$\tilde{y} = -\frac{at}{2\omega}\cos \omega t$$

et la solution générale de l'équation (9) est

$$y = \tilde{y} + Y = A \sin (\omega t + \varphi) - \frac{at}{2\omega}\cos \omega t.$$

On voit sur cette formule que comme dans le cas précédent la solution générale décrit la composition de deux mouvements oscillatoires mais de pulsations égales.

La présence du facteur t exprime que l'amplitude des oscillations croît indéfiniment avec t, c'est-à-dire que le solide effectuera au bout d'un intervalle de temps des oscillations d'amplitude très élevée même si l'amplitude a de la force extérieure est petite. Ce phénomène s'appelle *résonance*. Il conduit souvent à la destruction des systèmes oscillatoires.

Le cas des oscillations mécaniques de systèmes élastiques (telles les oscillations des ressorts de wagons, d'automobiles, etc.) se traite comme celui du solide suspendu à un ressort. On étudie de même les oscillations électriques, acoustiques et autres. Les équations différentielles linéaires sont capitales dans l'étude de tous ces problèmes.

Signalons en conclusion que la théorie des équations différentielles linéaires du second ordre se généralise *ad litteram* aux équations différentielles linéaires de tout ordre.

INDEX ALPHABÉTIQUE DES MATIÈRES

Abaissement de l'ordre 481
Accroissement de l'argument 98
— de la fonction 98
— partiel 317
Affixe 448
Aire élémentaire 344
Amplitude des oscillations 502
Angle de coordonnées 42
— polaire 45
Asymptotes d'une hyperbole 65, 164
Axes de coordonnées 16,41, 247
— de l'ellipse 62
— de l hyperbole 66
— imaginaire 448
— non transverse 66
— de la parabole 71
— polaire 45
— transverse 66

Borne(s) inférieure 18, 105
— d'intégration 363
— supérieure 18, 105

Carré scalaire 260
Centre d'une conique 74
— de l'ellipse 62
— de l'hyperbole 66
Champ(s) de directions 468
— laplacien (ou harmonique) 422
— non stationnaire 411
— potentiel 412
— scalaires 411
— solénoïdal 418
— stationnaire 411
— vectoriels 411
Circulation d'un champ 418
Coefficient (s) directeur 51
— de FOURIER 459
— de rigidité 501
Cofacteur 296
Condition(s) de CAUCHY 466
— initiale 466
— nécessaire d'extrémum 158, 337
— — d'inflexion 162
— suffisante d'extrémum 159, 337
— — d'inflexion 162
Cône du second ordre 288
Coniques 60
Conjugué complexe 449
Contour d'intégration 363
Convexité d'une courbe 160
Coordonnées 16
— cylindriques 391
— polaires 45, 391
— rectangulaires 41, 247
— sphériques 391
Cosinus directeurs 251
Courbe intégrale 465, 479
— du premier ordre 55
— du second ordre 60
Critère de d'ALEMBERT 429
— intégral 431
— de LEIBNIZ 434

Cylindre curviligne 345
— elliptique 272

Débit des puits 417
— des sources 417
Décomposition d'une fraction rationnelle 188
Définition axiomatique 15
δ-voisinage 17, 310
Demi-axe non transverse 67
— — transverse 67
— grand axe 62
— petit axe 62
Densité des sources (ou des puits) 417
Dérivée (s) finie 117
— d'une fonction 117,328
— infinie 117
— logarithmique 132
— partielle (s) 317
— — mixtes 330
— — secondes 330
— première 134
— seconde 134
Déterminant caractéristique 299
— fonctionnel 352
— principal 299
— de WRONSKI 486
Développement d'un déterminant 297
Diagonale non principale 294
— principale 292, 294
Diamètre d'un domaine 343
Différentielle 122, 325
— première 138, 333
— seconde 138,333
Directrice de l'ellipse 67
— de l'hyperbole 68
— de la parabole 70
Discontinuité de deuxième espèce 102
— de première espèce 102
Disque de convergence 454
Divergence 416
Domaine de convergence 436
— de définition 307
— d'intégration 344
— simplement connexe 476
— des valeurs 307
Droite orientée 15
— de régression 341

Ellipse 60
— imaginaire 75
Ensemble de définition 78
— des valeurs 78
Equation(s) canoniques du cône 288
— — de la droite 279
— — de l'ellipse 61
— — de l'ellipsoïde 283
— — de l'hyperbole 64
— — de l'hyperboloïde à deux nappes 285
— — — à une nappe 284
— — de la parabole 71
— — du paraboloïde elliptique 286
— — — hyperbolique 286
— caractéristique 493

Equation du centre de la conique 74
— complète 470
— d'une courbe 48
— aux différences totales 472
— différentielle du premier ordre 464
— — — — , linéaire 470
— — — — ,— homogène 470
— — — —, — non homogène 470
— — du second ordre 479
— — — —, linéaire 484
— — — —, — homogène 484
— — — —, — non homogène 484
— — à variables séparables 469
— de l'ellipse imaginaire 75
— générale de la droite 55, 278
— — du plan 274
— incomplète 55
— normale de la droite 58
— — du plan 276
— paramétrique 280
— des oscillations 501
— — forcées 501
— d'une surface 270
Excentricité 63, 67
Extrémum local 157

Facteur de normalisation 59, 277
Flux d'un vecteur 415
Fonction(s) 78
— bornée 79
— composée 112
— constante 78
— continue 97, 98, 99, 313
— — à droite 98
— — à gauche 98
— — par morceaux 102
— croissante 112
— décroissante 112
— définie paramétriquement 140
— dérivable 120
— de DIRICHLET 81
— élémentaires 81
— infiniment grande 93
— — petite 92
— intégrable 200
— irrationnelle 82
— linéaire 82
— linéairement dépendante 485
— — indépendante 485
— majorée 78
— minorée 78
— monotones 112
— mutuellement réciproques 115
— polynôme 82
— potentielle 412
— rationnelle 82
— réciproque 113
— scalaire 411
— strictement croissantes 112
— — décroissantes 112
— — monotones 113
— transcendantes 82
— uniformément continue 109
— vectorielle 411
Formule des accroissements finis 147
— — — généralisée 147
— du binôme de NEWTON 155
— de CAUCHY 147
— de changement de la variable d'intégration 183, 218
— de CRAMER 299
— d'EULER 455
— de GREEN 374
— d'intégration par parties 185, 220
— d'interpolation de LAGRANGE 171
— — de NEWTON 172
— de LAGRANGE 147
— de LEIBNIZ 136
Formule de MACLAURIN 154
— de MOIVRE 451
— de la moyenne 213
— de NEWTON-LEIBNIZ 216
— d'OSTROGRADSKI 405
— de PONCELET (des trapèzes) 240
— de récurrence 187
— de SIMPSON (ou des paraboles) 244
— de STOKES 409
— de TAYLOR 153, 334
Foyers 60, 70
Fraction rationnelle 82

Génératrice 272
Gradient d'une fonction 329
Grandeurs scalaires 248
— vectorielles 248
Graphique d'une fonction 79

Hyperbole 63
— conjuguée 67
— équilatère 67
Hyperboloïde à deux nappes 285
— à une nappe 284

Image 448
Inégalités larges 13
— strictes 13
Indéterminations 147
Infimum 18
Infiniment grand 93
— petit(s) 92, 312
— — équivalents 95
Intégrale (s) curvilignes 362
— — de première espèce 363
— — de seconde espèce 367
— définie 199
— double 344
— impropre de première espèce 233
— — de seconde espèce 236
— indéfinie 179
— itérée 347
— de surface 395
— — de première espèce 395
— — de seconde espèce 399
— tabulaires 182
— triple 387
Intégrant 179, 200
Intégration 179
— directe 182
Interpolation des fonctions 169
— d'HERMITE 172
Intervalles 17
— de convergence 438
Invariant 74
Inverse 13

Jacobien 352

Ligne de niveau 309
Limite à droite 86
— finie 29
— d'une fonction 82, 84, 87
— à gauche 86
— infinie 29
Longueur d'un arc 225
— d'un vecteur 248

Matrice colonne 291
— ligne 291
— d'une transformation 290
— unité 292
Maximum local strict 157
Mesure algébrique 16
Méthode des coefficients indéterminés 189

Méthode d'EULER 474
— de la fourchette 174
— d'intégration par substitution 182
— des moindres carrés 340
— des tangentes (de NEWTON) 174
Mineur d'un élément 296
Minimum local strict 157
Module d'un vecteur 248
Moments statiques 360

Nœuds d'interpolation 169
Nombre complexe 448
— — imaginaire 448
— — — pur 448
Normale à une surface 327

Octants 248
Opérateur de HAMILTON 420
— de LAPLACE 422
Ordonnée 41, 247
Orientation d'une surface 399
Oscillation 108

Parabole 70
Paraboloïde elliptique 285
— hyperbolique 286
— de révolution 286, 477
Paramètre 140
— de la parabole 70
Pente 51
Période de désintégration 478
Plan tangent 326
Points de discontinuité 101, 313
— — de deuxième espèce 102
— — de première espèce 102
— d'éventuel extrémum 159, 337
— d'extrémums locaux 336
— d'inflexion 161
— singulier 235, 466
— stationnaires 159
Pôle 45
Polynôme d'interpolation 169
Potentiel d'un champ 412
Primitive d'une fonction 178
Problème de CAUCHY 466
— d'interpolation 169
Produit 12
— scalaire 258
— mixte 268
— vectoriel 262
Projeté d'un vecteur 249
Prolongement périodique 459
Pulsation 502

Quotient 14

Rapport algébrique 44
Rayon de convergence 438, 454
— polaire 45
— vecteur 60
Règle de l'HÔPITAL 149
— de SARRUS 294
Relation de CHASLES 16
Résonance 505
Reste de l'interpolation 173
— de LAGRANGE 153
— de PEANO 154
Rotation des axes 46
Rotationnel 418

Sécante 118
Secteur curviligne 223
Segment orienté 15
Sens direct (ou positif) 370
— rétrograde (ou négatif) 370
Série(s) absolument convergente 435, 453
— alternées 433
Série(s) convergente 452
— entière 436, 453
— de FOURIER 459
— harmonique alternée 434
— de MACLAURIN 442
— numérique 423
— semi-convergentes 436
— à termes complexes 448
— — positifs 428
— trigonométrique 456
Solution d'une équation 465
— générale 466, 480
— particulière 466, 480
Somme(s) de DARBOUX 202
— intégrale 199, 343
— partielle d'une série 423
— d'une série 423, 452
Sommets de l'ellipse 62
— de l'hyperbole 66
— de la parabole 71
Spirale d'ARCHIMÈDE 49
Suite bornée 24
— convergente 28, 310
— croissante 37
— décroissante 37
— divergente 28
— illimitée 25
— infiniment grande 25
— — petite 25
— majorée 24
— minorée 24
— monotone 37
— numérique 23
— strictement croissante 37
— — décroissante 37
— — monotone 37
— de segments emboîtés 40
Support d'un vecteur 248
Surface(s) différentiable 395
— non orientables 399
— orientables 399
— du second ordre 283

Théorème(s) d'ABEL 437
— de BOLZANO-CAUCHY 103, 104
— de CANTOR 109
— de CAUCHY 146, 456, 479
— de FERMAT 143
— de l'HÔPITAL 148
— de LAGRANGE 145
— de la moyenne 213, 346
— des projections 256
— de TAYLOR 152
— de WEIERSTRASS 106, 107
Translation 46
Transformation identique 293
Trièdre direct 262
— effectif 262
— rétrograde 262

Valeur absolue 19
— principale 450
Variable dépendante 78
— indépendante 78
— intermédiaire 112
Vecteur(s) coplanaires 262
— directeur 278
— équipollents 248
— libre (s) 248
— — égaux 248
— lié 248
— normal 274
— potentiel 412
Vitesse instantanée 119
— moyenne 119

Wronskien 486

TABLE DES MATIÈRES

Préface 5
Avant-propos 6
Introduction 7

Première partie

ANALYSE MATHÉMATIQUE
DES FONCTIONS D'UNE VARIABLE

Chapitre premier. NOMBRES RÉELS 11

§ 1. Ensembles. Notations. Symboles logiques 11
§ 2. Les nombres réels et leurs propriétés fondamentales . . . 12
§ 3. Représentation géométrique des nombres réels 15
§ 4. Bornes des ensembles numériques 18
§ 5. Valeur absolue d'un nombre 19
§ 6. Raisonnement par récurrence 21

Chapitre II. LIMITE D'UNE SUITE 23

§ 1. Suites numériques 23
§ 2. Suites convergentes 28
§ 3. Suites monotones 37
§ 4. Théorème des segments emboîtés 39

Chapitre III. GÉOMÉTRIE ANALYTIQUE DANS LE PLAN . . 41

§ 1. Système de coordonnées rectangulaires 41
§ 2. Problèmes élémentaires de géométrie analytique dans le plan 42
§ 3. Coordonnées polaires 45
§ 4. Changement des coordonnées rectangulaires 46
§ 5. Equation d'une courbe dans le plan 48
§ 6. Courbes du premier ordre 50
§ 7. Courbes du second ordre 60
§ 8. Equation générale des coniques 72

Chapitre IV. FONCTIONS D'UNE VARIABLE 78

§ 1. Notion de fonction 78
§ 2. Limite d'une fonction 82
§ 3. Théorèmes relatifs aux limites des fonctions 88
§ 4. Deux limites remarquables 89
§ 5. Infiniment petits et infiniment grands 92

§ 6. Comparaison des infiniment petits et des infiniment grands 95
§ 7. Notion de continuité d'une fonction 97
§ 8. Continuité de certaines fonctions élémentaires 99
§ 9. Classification des points de discontinuité d'une fonction 101
§ 10. Propriétés fondamentales des fonctions continues 102
§ 11. Notion de fonction composée 112
§ 12. Notion de fonction réciproque 112

Chapitre V. DÉRIVATION 117

§ 1. Notion de dérivée 117
§ 2. Notion de dérivabilité d'une fonction 120
§ 3. Notion de différentielle 122
§ 4. Règles de dérivation d'une somme, d'une différence, d'un produit et d'un quotient 124
§ 5. Calcul des dérivées des fonctions constante, puissance, circulaires et logarithme 125
§ 6. Théorème de la dérivée de la fonction réciproque 128
§ 7. Calcul des dérivées de la fonction exponentielle et des fonctions circulaires réciproques 129
§ 8. Règle de dérivation d'une fonction composée 130
§ 9. Dérivée logarithmique. Dérivée d'une fonction puissance à exposant réel quelconque. Table des dérivées des fonctions élémentaires 132
§ 10. Dérivées et différentielles supérieures 134
§ 11. Définition paramétrique d'une fonction et dérivation 139

Chapitre VI. APPLICATION DU CALCUL DIFFÉRENTIEL À L'ÉTUDE DES FONCTIONS 143

§ 1. Théorèmes fondamentaux du calcul différentiel 143
§ 2. Levée des indéterminations. Règle de l'Hôpital 147
§ 3. Formule de Taylor 152
§ 4. Etude des fonctions et construction de leurs courbes représentatives 157
§ 5. Interpolation des fonctions 169
§ 6. Méthode de calcul approché des racines d'une équation 174

Chapitre VII. INTÉGRALE INDÉFINIE 178

§ 1. Primitive et intégrale indéfinie 178
§ 2. Propriétés fondamentales de l'intégrale indéfinie 180
§ 3. Table des principales intégrales 181
§ 4. Principales méthodes d'intégration 182
§ 5. Intégration des fonctions rationnelles 187
§ 6. Intégration des fonctions irrationnelles et transcendantes 192

Chapitre VIII. INTÉGRALE DÉFINIE 199

§ 1. Définition de l'intégrale définie 199
§ 2. Conditions d'existence de l'intégrale définie 201
§ 3. Intégrabilité des fonctions continues et de certaines fonctions discontinues 207
§ 4. Propriétés fondamentales de l'intégrale définie 208
§ 5. Estimations des intégrales. Formule de la moyenne . . . 211
§ 6. Intégrale à borne supérieure variable 214
§ 7. Formule de Newton-Leibniz 215

§ 8. Changement de variable dans l'intégrale définie 217
§ 9. Formule d'intégration par parties dans l'intégrale définie 219
§ 10. Quelques applications physiques et géométriques de l'intégrale définie 220
§ 11. Intégrales impropres 233
§ 12. Calcul approché des intégrales définies 240

Deuxième partie

ANALYSE MATHÉMATIQUE DES FONCTIONS DE PLUSIEURS VARIABLES

Chapitre IX. GÉOMÉTRIE ANALYTIQUE DANS L'ESPACE . . 247

§ 1. Système de coordonnées rectangulaires dans l'espace . . . 247
§ 2. Notion de vecteur 248
§ 3. Opérations linéaires sur les vecteurs et leurs propriétés fondamentales 252
§ 4. Théorèmes des projections 256
§ 5. Décomposition d'un vecteur suivant une base 257
§ 6. Produit scalaire des vecteurs 258
§ 7. Produit vectoriel 262
§ 8. Produit mixte de trois vecteurs 268
§ 9. Equations d'une surface et d'une ligne 270
§ 10. Equation d'un cylindre 272
§ 11. Equations du plan 273
§ 12. Equations de la droite 278
§ 13. Position relative d'une droite et d'un plan 282
§ 14. Quadriques 283

Chapitre X. ÉLÉMENTS D'ALGÈBRE SUPÉRIEURE . . 289

§ 1. Matrices 289
§ 2. Déterminants 293
§ 3. Système de trois équations du premier degré à trois inconnues 298
§ 4. Ecriture matricielle d'un système d'équations linéaires. Notion de matrice inverse 303

Chapitre XI. FONCTIONS DE PLUSIEURS VARIABLES: NOTION, LIMITE ET CONTINUITÉ 307

§ 1. Notion de fonction de plusieurs variables 307
§ 2. Représentation géométrique des fonctions de deux variables 309
§ 3. Limite d'une fonction de deux variables 310
§ 4. Continuité d'une fonction de deux variables 313

Chapitre XII. DÉRIVÉES PARTIELLES ET DÉRIVABILITÉ DES FONCTIONS DE PLUSIEURS VARIABLES 317

§ 1. Dérivées partielles 317
§ 2. Notion de dérivabilité d'une fonction 318
§ 3. Dérivées de fonctions composées 322
§ 4. Différentielle d'une fonction 325
§ 5. Dérivée suivant une direction. Gradient 327
§ 6. Dérivées partielles et différentielles supérieures 330

§ 7. Formulle de Taylor pour les fonctions de deux variables . . 334
§ 8. Extrémums d'une fonction de deux variables 336
§ 9. Méthode des moindres carrés 340

Chapitre XIII. INTÉGRATION 343

§ 1. Intégrales doubles 343
§ 2. Réduction de l'intégrale double à l'intégrale itérée . . . 346
§ 3. Changement de variables dans l'intégrale double 351
§ 4. Quelques applications géométriques et physiques des intégrales doubles 354
§ 5. Intégrales curvilignes 362
§ 6. Formule de Green 373
§ 7. Conditions d'indépendance de l'intégrale curviligne par rapport au chemin d'intégration 376
§ 8. Intégration des différentielles totales 380
§ 9. Quelques applications des intégrales curvilignes de seconde espèce 384
§ 10. Intégrales triples 387
§ 11. Intégrales de surface 394
§ 12. Formule d'Ostrogradski 404
§ 13. Formule de Stokes 407
§ 14. Champs scalaire et vectoriel 410

Troisième partie

SÉRIES, ÉQUATIONS DIFFÉRENTIELLES

Chapitre XIV. SÉRIES 423

§ 1. Notion de série numérique 423
§ 2. Séries à termes positifs 428
§ 3. Séries alternées 433
§ 4. Convergence absolue et semi-convergence des séries 435
§ 5. Séries entières 436
§ 6. Séries à termes complexes 448
§ 7. Séries de Fourier 456

Chapitre XV. ÉQUATIONS DIFFÉRENTIELLES ORDINAIRES 464

§ 1. Equations différentielles du premier ordre 464
§ 2. Equations différentielles du second ordre 479
§ 3. Equations différentielles linéaires du second ordre 484
§ 4. Equations différentielles linéaires du second ordre à coefficients constants 492
§ 5. Application des équations différentielles linéaires à l'étude des phénomènes vibratoires 500

Index alphabétique des matières 506